

Polarization Spectroscopy: Principles, Theory, Techniques and Applications

Biological Applications

Contents

Articles

Copyright@2009 by Bci2 1

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version ;published by the Free Software Foundation, with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in ;the section entitled "GNU Free Documentation License". 2

Edited by Bci2, with the contributors listed after the Common Use License. 3

Spectroscopy--An Introduction 4

Spectroscopy 4

Fourier transform spectroscopy 11

Spectroscopy Theory 15

Quantum mechanics 15

Quantum field theory 30

Algebraic quantum field theory 41

Local quantum field theory 42

Algebraic logic 43

Quantum logic 46

Quantum computer 53

Quantum chemistry 62

Density functional theory 66

Birefringence 73

Polarization spectroscopy 79

Polarized IR Spectroscopy 79

Circular dichroism 85

Vibrational circular dichroism 91

Optical rotatory dispersion 101

Raman spectroscopy 101

Coherent anti-Stokes Raman spectroscopy 107

Raman Microscopy	110
Imaging spectroscopy	110
Chemical imaging	114
Spin polarization	121
Polarized Neutron Spectroscopy	122
Polarized Muon Spectroscopy	124
Time-resolved spectroscopy	126
Terahertz spectroscopy	127
Applied spectroscopy	128
Amino acids	131
Proteins	144
Protein structure	159
Protein folding	167
Protein dynamics	174
Nucleic Acids	189
DNA	192
Molecular models of DNA	216
DNA structure	224
DNA Dynamics	232
Interactomics	239

References

Article Sources and Contributors	242
Image Sources, Licenses and Contributors	247

Article Licenses

License	250
---------	-----

Copyright@2009 by Bci2

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version ;published by the Free Software Foundation, with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in ;the section entitled "GNU Free Documentation License".

Edited by Bci2, with the contributors listed
after the Common Use License.

Spectroscopy--An Introduction

Spectroscopy

Spectroscopy was originally the study of the interaction between radiation and matter as a function of wavelength (λ). In fact, historically, spectroscopy referred to the use of visible light dispersed according to its wavelength, e.g. by a prism. Later the concept was expanded greatly to comprise any measurement of a quantity as a function of either wavelength or frequency. Thus it also can refer to a response to an alternating field or varying frequency (ν). A further extension of the scope of the definition added energy (E) as a variable, once the very close relationship $E = h\nu$ for photons was realized (h is the Planck constant). A plot of the response as a function of wavelength—or more commonly frequency—is referred to as a spectrum; see also spectral linewidth.

Spectrometry is the spectroscopic technique used to assess the concentration or amount of a given chemical (atomic, molecular, or ionic) species. In this case, the instrument that performs such measurements is a spectrometer, spectrophotometer, or spectrograph.

Spectroscopy/spectrometry is often used in physical and analytical chemistry for the identification of substances through the spectrum emitted from or absorbed by them.

Spectroscopy/spectrometry is also heavily used in astronomy and remote sensing. Most large telescopes have spectrometers, which are used either to measure the chemical composition and physical properties of astronomical objects or to measure their velocities from the Doppler shift of their spectral lines.

Classification of methods

Nature of excitation measured

The type of spectroscopy depends on the physical quantity measured. Normally, the quantity that is measured is an intensity, either of energy absorbed or produced.

- Electromagnetic spectroscopy involves interactions of matter with electromagnetic radiation, such as light.
- Electron spectroscopy involves interactions with electron beams. Auger spectroscopy involves inducing the Auger effect with an electron beam. In this case the measurement typically involves the kinetic energy of the electron as variable.
- Acoustic spectroscopy involves the frequency of sound.
- Dielectric spectroscopy involves the frequency of an external electrical field
- Mechanical spectroscopy involves the frequency of an external mechanical stress, e.g. a torsion applied to a piece of material.

Measurement process

Most spectroscopic methods are differentiated as either atomic or molecular based on whether or not they apply to atoms or molecules. Along with that distinction, they can be classified on the nature of their interaction:

- Absorption spectroscopy uses the range of the electromagnetic spectra in which a substance absorbs. This includes atomic absorption spectroscopy and various molecular techniques, such as infrared, ultraviolet-visible and microwave spectroscopy.
 - Emission spectroscopy uses the range of electromagnetic spectra in which a substance radiates (emits). The substance first must absorb energy. This energy can be from a variety of sources, which determines the name of
-

the subsequent emission, like luminescence. Molecular luminescence techniques include spectrofluorimetry.

- Scattering spectroscopy measures the amount of light that a substance scatters at certain wavelengths, incident angles, and polarization angles. One of the most useful applications of light scattering spectroscopy is Raman spectroscopy.

Common types

Absorption

Absorption spectroscopy is a technique in which the power of a beam of light measured before and after interaction with a sample is compared. Specific absorption techniques tend to be referred to by the wavelength of radiation measured such as ultraviolet, infrared or microwave absorption spectroscopy. Absorption occurs when the energy of the photons matches the energy difference between two states of the material.

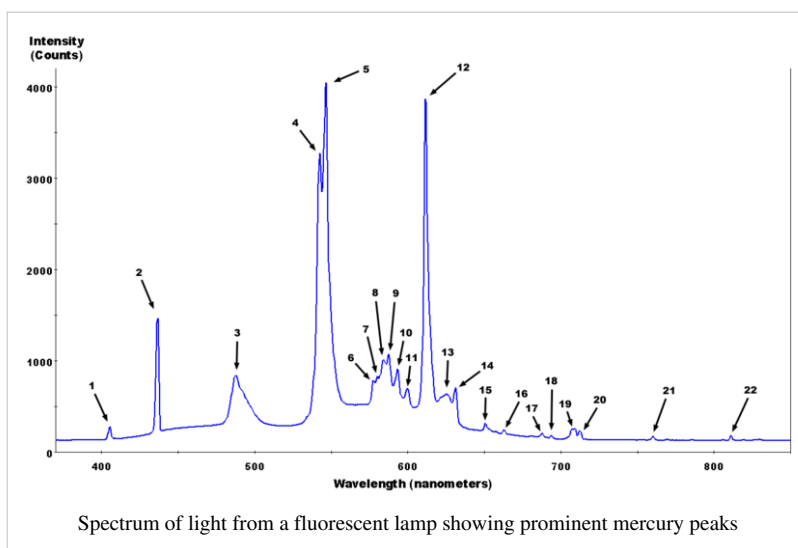
Fluorescence

Fluorescence spectroscopy uses higher energy photons to excite a sample, which will then emit lower energy photons. This technique has become popular for its biochemical and medical applications, and can be used for confocal microscopy, fluorescence resonance energy transfer, and fluorescence lifetime imaging.

X-ray

When X-rays of sufficient frequency (energy) interact with a substance, inner shell electrons in the atom are excited to outer empty orbitals, or they may be removed completely, ionizing the atom. The inner shell "hole" will then be filled by electrons from outer orbitals. The energy available in this de-excitation process is emitted as radiation (fluorescence) or will remove other less-bound electrons from the atom (Auger effect). The absorption or emission frequencies (energies) are characteristic of the specific atom. In addition, for a specific atom small frequency (energy) variations occur which are characteristic of the chemical bonding. With a suitable apparatus, these characteristic X-ray frequencies or Auger electron energies can be measured. X-ray absorption and emission spectroscopy is used in chemistry and material sciences to determine elemental composition and chemical bonding.

X-ray crystallography is a scattering process; crystalline materials scatter X-rays at well-defined angles. If the wavelength of the incident X-rays is known, this allows calculation of the distances between planes of atoms within the crystal. The intensities of the scattered X-rays give information about the atomic positions and allow the arrangement of the atoms within the crystal structure to be calculated. However, the X-ray light is then not dispersed according to its wavelength, which is set at a given value, and X-ray diffraction is thus not a spectroscopy.



Flame

Liquid solution samples are aspirated into a burner or nebulizer/burner combination, desolvated, atomized, and sometimes excited to a higher energy electronic state. The use of a flame during analysis requires fuel and oxidant, typically in the form of gases. Common fuel gases used are acetylene (ethyne) or hydrogen. Common oxidant gases used are oxygen, air, or nitrous oxide. These methods are often capable of analyzing metallic element analytes in the part per million, billion, or possibly lower concentration ranges. Light detectors are needed to detect light with the analysis information coming from the flame.

- **Atomic Emission Spectroscopy** - This method uses flame excitation; atoms are excited from the heat of the flame to emit light. This method commonly uses a total consumption burner with a round burning outlet. A higher temperature flame than atomic absorption spectroscopy (AA) is typically used to produce excitation of analyte atoms. Since analyte atoms are excited by the heat of the flame, no special elemental lamps to shine into the flame are needed. A high resolution polychromator can be used to produce an emission intensity vs. wavelength spectrum over a range of wavelengths showing multiple element excitation lines, meaning multiple elements can be detected in one run. Alternatively, a monochromator can be set at one wavelength to concentrate on analysis of a single element at a certain emission line. Plasma emission spectroscopy is a more modern version of this method. See Flame emission spectroscopy for more details.
- **Atomic absorption spectroscopy** (often called AA) - This method commonly uses a pre-burner nebulizer (or nebulizing chamber) to create a sample mist and a slot-shaped burner which gives a longer pathlength flame. The temperature of the flame is low enough that the flame itself does not excite sample atoms from their ground state. The nebulizer and flame are used to desolvate and atomize the sample, but the excitation of the analyte atoms is done by the use of lamps shining through the flame at various wavelengths for each type of analyte. In AA, the amount of light absorbed after going through the flame determines the amount of analyte in the sample. A graphite furnace for heating the sample to desolvate and atomize is commonly used for greater sensitivity. The graphite furnace method can also analyze some solid or slurry samples. Because of its good sensitivity and selectivity, it is still a commonly used method of analysis for certain trace elements in aqueous (and other liquid) samples.
- **Atomic Fluorescence Spectroscopy** - This method commonly uses a burner with a round burning outlet. The flame is used to solvate and atomize the sample, but a lamp shines light at a specific wavelength into the flame to excite the analyte atoms in the flame. The atoms of certain elements can then fluoresce emitting light in a different direction. The intensity of this fluorescing light is used for quantifying the amount of analyte element in the sample. A graphite furnace can also be used for atomic fluorescence spectroscopy. This method is not as commonly used as atomic absorption or plasma emission spectroscopy.

Plasma Emission Spectroscopy In some ways similar to flame atomic emission spectroscopy, it has largely replaced it.

- Direct-current plasma (DCP)

A direct-current plasma (DCP) is created by an electrical discharge between two electrodes. A plasma support gas is necessary, and Ar is common. Samples can be deposited on one of the electrodes, or if conducting can make up one electrode.

- Glow discharge-optical emission spectrometry (GD-OES)
- Inductively coupled plasma-atomic emission spectrometry (ICP-AES)
- Laser Induced Breakdown Spectroscopy (LIBS) (LIBS), also called Laser-induced plasma spectrometry (LIPS)
- Microwave-induced plasma (MIP)

Spark or arc (emission) spectroscopy - is used for the analysis of metallic elements in solid samples. For non-conductive materials, a sample is ground with graphite powder to make it conductive. In traditional arc spectroscopy methods, a sample of the solid was commonly ground up and destroyed during analysis. An electric arc

or spark is passed through the sample, heating the sample to a high temperature to excite the atoms in it. The excited analyte atoms glow emitting light at various wavelengths which could be detected by common spectroscopic methods. Since the conditions producing the arc emission typically are not controlled quantitatively, the analysis for the elements is qualitative. Nowadays, the spark sources with controlled discharges under an argon atmosphere allow that this method can be considered eminently quantitative, and its use is widely expanded worldwide through production control laboratories of foundries and steel mills.

Visible

Many atoms emit or absorb visible light. In order to obtain a fine line spectrum, the atoms must be in a gas phase. This means that the substance has to be vaporised. The spectrum is studied in absorption or emission. Visible absorption spectroscopy is often combined with UV absorption spectroscopy in UV/Vis spectroscopy. Although this form may be uncommon as the human eye is a similar indicator, it still proves useful when distinguishing colours.

Ultraviolet

All atoms absorb in the Ultraviolet (UV) region because these photons are energetic enough to excite outer electrons. If the frequency is high enough, photoionization takes place. UV spectroscopy is also used in quantifying protein and DNA concentration as well as the ratio of protein to DNA concentration in a solution. Several amino acids usually found in protein, such as tryptophan, absorb light in the 280 nm range and DNA absorbs light in the 260 nm range. For this reason, the ratio of 260/280 nm absorbance is a good general indicator of the relative purity of a solution in terms of these two macromolecules. Reasonable estimates of protein or DNA concentration can also be made this way using Beer's law.

Infrared

Infrared spectroscopy offers the possibility to measure different types of inter atomic bond vibrations at different frequencies. Especially in organic chemistry the analysis of IR absorption spectra shows what type of bonds are present in the sample. It is also an important method for analysing polymers and constituents like fillers, pigments and plasticizers.

Near Infrared (NIR)

The near infrared NIR range, immediately beyond the visible wavelength range, is especially important for practical applications because of the much greater penetration depth of NIR radiation into the sample than in the case of mid IR spectroscopy range. This allows also large samples to be measured in each scan by NIR spectroscopy, and is currently employed for many practical applications such as: rapid grain analysis, medical diagnosis pharmaceuticals/medicines^[1], biotechnology, genomics analysis, proteomic analysis, interactomics research, inline textile monitoring, food analysis and chemical imaging/hyperspectral imaging of intact organisms^{[2] [3] [4]}, plastics, textiles, insect detection, forensic lab application, crime detection, various military applications, and so on.

Raman

Raman spectroscopy uses the inelastic scattering of light to analyse vibrational and rotational modes of molecules. The resulting 'fingerprints' are an aid to analysis.

Coherent anti-Stokes Raman spectroscopy (CARS)

CARS is a recent technique that has high sensitivity and powerful applications for *in vivo* spectroscopy and imaging^[5].

Nuclear magnetic resonance

Nuclear magnetic resonance spectroscopy analyzes the magnetic properties of certain atomic nuclei to determine different electronic local environments of hydrogen, carbon, or other atoms in an organic compound or other compound. This is used to help determine the structure of the compound.

Mössbauer

Transmission or conversion-electron (CEMS) modes of Mössbauer spectroscopy probe the properties of specific isotope nuclei in different atomic environments by analyzing the resonant absorption of characteristic energy gamma-rays known as the Mössbauer effect.

Other types

There are many different types of materials analysis techniques under the broad heading of "spectroscopy", utilizing a wide variety of different approaches to probing material properties, such as absorbance, reflection, emission, scattering, thermal conductivity, and refractive index.

- Acoustic spectroscopy
 - Auger spectroscopy is a method used to study surfaces of materials on a micro-scale. It is often used in connection with electron microscopy.
 - Cavity ring down spectroscopy
 - Circular Dichroism spectroscopy
 - Deep-level transient spectroscopy measures concentration and analyzes parameters of electrically active defects in semiconducting materials
 - Dielectric spectroscopy
 - Dual polarisation interferometry measures the real and imaginary components of the complex refractive index
 - Force spectroscopy
 - Fourier transform spectroscopy is an efficient method for processing spectra data obtained using interferometers. Nearly all infrared spectroscopy techniques (such as FTIR) and nuclear magnetic resonance (NMR) are based on Fourier transforms.
 - Fourier transform infrared spectroscopy (FTIR)
 - Hadron spectroscopy studies the energy/mass spectrum of hadrons according to spin, parity, and other particle properties. Baryon spectroscopy and meson spectroscopy are both types of hadron spectroscopy.
 - Inelastic electron tunneling spectroscopy (IETS) uses the changes in current due to inelastic electron-vibration interaction at specific energies which can also measure optically forbidden transitions.
 - Inelastic neutron scattering is similar to Raman spectroscopy, but uses neutrons instead of photons.
 - Laser spectroscopy uses lasers^[6] and other types of coherent emission sources, such as optical parametric oscillators,^[7] for selective excitation of atomic or molecular species.
 - Ultra fast laser spectroscopy
-

- Mechanical spectroscopy involves interactions with macroscopic vibrations, such as phonons. An example is acoustic spectroscopy, involving sound waves.
- Neutron spin echo spectroscopy measures internal dynamics in proteins and other soft matter systems
- Nuclear magnetic resonance (NMR)
- Photoacoustic spectroscopy measures the sound waves produced upon the absorption of radiation.
- Photothermal spectroscopy measures heat evolved upon absorption of radiation.
- Raman optical activity spectroscopy exploits Raman scattering and optical activity effects to reveal detailed information on chiral centers in molecules.
- Terahertz spectroscopy uses wavelengths above infrared spectroscopy and below microwave or millimeter wave measurements.
- Time-resolved spectroscopy is the spectroscopy of matter in situations where the properties are changing with time.
- Thermal infrared spectroscopy measures thermal radiation emitted from materials and surfaces and is used to determine the type of bonds present in a sample as well as their lattice environment. The techniques are widely used by organic chemists, mineralogists, and planetary scientists.

Background subtraction

Background subtraction is a term typically used in spectroscopy when one explains the process of acquiring a background radiation level (or ambient radiation level) and then makes an algorithmic adjustment to the data to obtain qualitative information about any deviations from the background, even when they are an order of magnitude less decipherable than the background itself.

Background subtraction can affect a number of statistical calculations (Continuum, Compton, Bremsstrahlung) leading to improved overall system performance.

Applications

- Estimate weathered wood exposure times using Near infrared spectroscopy.^[8]
- Cure monitoring of composites using Optical fibers

See also

- Absorption cross section
 - Applied spectroscopy
 - Astronomical spectroscopy
 - Atomic spectroscopy
 - Nuclear magnetic resonance
 - 2D-FT NMRI and Spectroscopy
 - 2D correlation analysis
 - Near infrared spectroscopy
 - Coherent spectroscopy
 - Cold vapour atomic fluorescence spectroscopy
 - Deep-level transient spectroscopy
 - EPR spectroscopy
 - Gamma spectroscopy
 - Kelvin probe force microscope
 - Metamerism (color)
 - Rigid rotor
-

- Rotational spectroscopy
- Saturated spectroscopy
- Scanning tunneling spectroscopy
- Scattering theory
- Spectral power distributions
- Spectral reflectance
- Spectrophotometry
- Spectroscopic notation
- Spectrum analysis
- The Unscrambler (CAMO Software)
- Vibrational spectroscopy
- Vibrational circular dichroism spectroscopy
- Robert Bunsen
- Gustav Kirchhoff
- Joseph von Fraunhofer

External links

- Spectroscopy links ^[9] at the Open Directory Project
- Amateur spectroscopy links ^[10] at the Open Directory Project
- Timeline of Spectroscopy ^[11]
- Chemometric Analysis for Spectroscopy ^[12]
- The Science of Spectroscopy ^[13] - supported by NASA, includes OpenSpectrum, a Wiki-based learning tool for spectroscopy that anyone can edit
- A Short Study of the Characteristics of two Lab Spectroscopes ^[14]
- NIST government spectroscopy data ^[15]
- Potentiodynamic Electrochemical Impedance Spectroscopy ^[16]

References

- [1] J. Dubois, G. Sando, E. N. Lewis, Near-Infrared Chemical Imaging, A Valuable Tool for the Pharmaceutical Industry, G.I.T. Laboratory Journal Europe, No. 1-2, 2007
- [2] http://www.malvern.com/LabEng/products/sdi/bibliography/sdi_bibliography.htm E. N. Lewis, E. Lee and L. H. Kidder, Combining Imaging and Spectroscopy: Solving Problems with Near-Infrared Chemical Imaging. Microscopy Today, Volume 12, No. 6, 11/2004.
- [3] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis*, D. Luthria, Editor pp.241-273, AOCs Press., Champaign, IL.
- [4] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann and S. S. Korban, q-bio/0407006 (July 2004) (<http://arxiv.org/abs/q-bio/0407006>)
- [5] C.L. Evans and X.S. Xie.2008. Coherent Anti-Stokes Raman Scattering Microscopy: Chemical Imaging for Biology and Medicine., doi:10.1146/annurev.anchem.1.031207.112754 *Annual Review of Analytical Chemistry*, **1**: 883-909.
- [6] W. Demtröder, *Laser Spectroscopy*, 3rd Ed. (Springer, 2003).
- [7] F. J. Duarte (Ed.), *Tunable Laser Applications*, 2nd Ed. (CRC, 2009) Chapter 2. (<http://www.opticsjournal.com/tla.htm>)
- [8] "Using NIR Spectroscopy to Predict Weathered Wood Exposure Times" (http://www.fpl.fs.fed.us/documnts/pdf2006/fpl_2006_wang002.pdf). .
- [9] <http://www.dmoz.org/Science/Physics/Optics/Spectroscopy/>
- [10] <http://www.dmoz.org/Science/Astronomy/Amateur/Spectroscopy/>
- [11] <http://spectroscopyonline.findanalytichem.com/spectroscopy/article/articleDetail.jsp?id=381944&sk=&date=&pageID=8>
- [12] <http://www.laboratoryequipment.com/article-chemometric-analysis-for-spectroscopy.aspx>
- [13] <http://www.scienceofspectroscopy.info>
- [14] <http://ioannis.virtualcomposer2000.com/spectroscopy/>
- [15] <http://physics.nist.gov/Pubs/AtSpec/index.html>
- [16] <http://www.abc.chemistry.bsu.by/vi/>

Fourier transform spectroscopy

Fourier transform spectroscopy is a measurement technique whereby spectra are collected based on measurements of the coherence of a radiative source, using time-domain or space-domain measurements of the electromagnetic radiation or other type of radiation. It can be applied to a variety of types of spectroscopy including optical spectroscopy, infrared spectroscopy (FT IR, FT-NIRS), Fourier transform (FT) nuclear magnetic resonance^[1], mass spectrometry and electron spin resonance spectroscopy. There are several methods for measuring the temporal coherence of the light, including the continuous wave *Michelson* or *Fourier transform* spectrometer and the pulsed Fourier transform spectrograph (which is more sensitive and has a much shorter sampling time than conventional spectroscopic techniques, but is only applicable in a laboratory environment).

Conceptual introduction

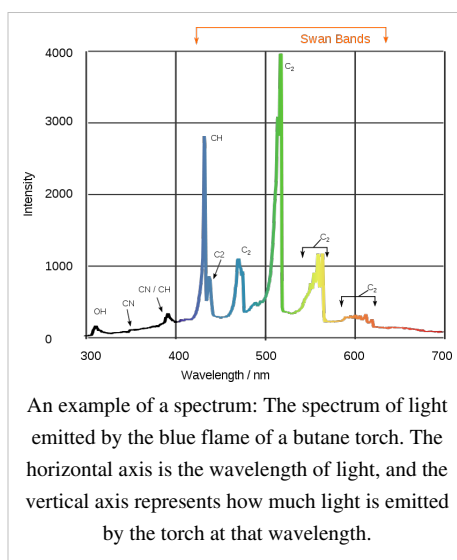
Measuring an emission spectrum

One of the most basic tasks in spectroscopy is to characterize the spectrum of a light source: How much light is emitted at each different wavelength. The most straightforward way to measure a spectrum is to pass the light through a monochromator, an instrument that blocks all of the light *except* the light at a certain wavelength (the un-blocked wavelength is set by a knob on the monochromator). Then the intensity of this remaining (single-wavelength) light is measured. The measured intensity directly indicates how much light is emitted at that wavelength. By varying the monochromator's wavelength setting, the full spectrum can be measured. This simple scheme in fact describes how *some* spectrometers work.

Fourier transform spectroscopy is a less intuitive way to get the same information. Rather than allowing only one wavelength at a time to pass through to the detector, this technique lets through a beam containing many different wavelengths of light at once, and measures the *total* beam intensity. Next, the beam is modified to contain a *different* combination of wavelengths, giving a second data point. This process is repeated many times. Afterwards, a computer takes all this data and works backwards to infer how much light there is at each wavelength.

To be more specific, between the light source and the detector, there is a certain configuration of mirrors that allows some wavelengths to pass through but blocks others (due to wave interference). The beam is modified for each new data point by moving one of the mirrors; this changes the set of wavelengths that can pass through.

As mentioned, computer processing is required to turn the raw data (light intensity for each mirror position) into the desired result (light intensity for each wavelength). The processing required turns out to be a common algorithm called the Fourier transform (hence the name, "Fourier transform spectroscopy"). The raw data is sometimes called an "interferogram".



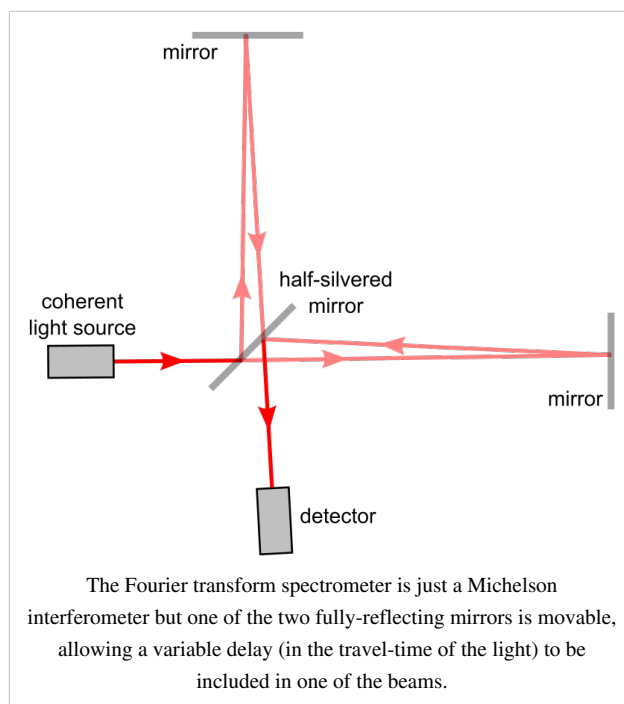
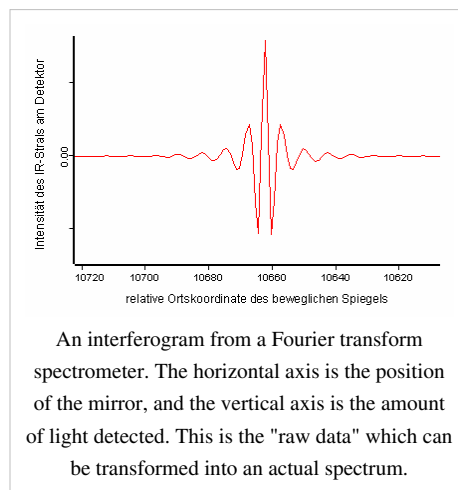
Measuring an absorption spectrum

The method of Fourier transform spectroscopy can also be used for absorption spectroscopy. The primary example is "FTIR Spectroscopy", a common technique in chemistry. In general, the goal of absorption spectroscopy is to measure how well a sample absorbs or transmits light at each different wavelength. However, any technique for *emission* spectroscopy can also be used for *absorption* spectroscopy as follows: Assume you have a working spectrometer that can measure the spectrum of any light source shining into it (as described above). Shine a broadband light source into the spectrometer, then shine the same light source *through the sample* into the same spectrometer. Comparing the two spectra, it will be obvious which wavelengths get absorbed by the sample and which wavelengths can pass right through the sample. (More precisely, the spectrum with the sample, divided by the "background" spectrum without the sample, equals the fraction of light that the sample can transmit at each wavelength.)

Accordingly, the technique of "Fourier transform spectroscopy" can be used both for measuring emission spectra (for example, the emission spectrum of a star), *and* absorption spectra (for example, the absorption spectrum of a glass of liquid).

Continuous wave *Michelson* or *Fourier transform* spectrograph

The Michelson spectrograph is similar to the instrument used in the Michelson-Morley experiment. Light from the source is split into two beams by a half-silvered mirror, one is reflected off a fixed mirror and one off a moving mirror which introduces a time delay -- the Fourier transform spectrometer is just a Michelson interferometer with a movable mirror. The beams interfere, allowing the temporal coherence of the light to be measured at each different time delay setting, effectively converting the time domain into a spatial coordinate. By making measurements of the signal at many discrete positions of the moving mirror, the spectrum can be reconstructed using a Fourier transform of the temporal coherence of the light. Michelson spectrographs are capable of very high spectral resolution observations of very bright sources. The Michelson or Fourier transform spectrograph was popular for infra-red applications at a time when infra-red astronomy only had single pixel detectors. Imaging Michelson spectrometers are a possibility, but in general have been supplanted by imaging Fabry-Perot instruments which are easier to construct.



Extracting the spectrum

The intensity as a function of the path length difference in the interferometer p and wavenumber $\tilde{\nu} = 1/\lambda$ is ^[2]

$$I(p, \tilde{\nu}) = I(\tilde{\nu})[1 + \cos(2\pi\tilde{\nu}p)],$$

where $I(\tilde{\nu})$ is the spectrum to be determined. Note that it is not necessary for $I(\tilde{\nu})$ to be modulated by the sample before the interferometer. In fact, most FTIR spectrometers place the sample after the interferometer in the optical path. The total intensity at the detector is

$$I(p) = \int_0^\infty I(p, \tilde{\nu}) d\tilde{\nu} = \int_0^\infty I(\tilde{\nu})[1 + \cos(2\pi\tilde{\nu}p)] d\tilde{\nu}.$$

This is just a Fourier cosine transform. The inverse gives us our desired result in terms of the measured quantity $I(p)$:

$$I(\tilde{\nu}) = 4 \int_0^\infty [I(p) - \frac{1}{2}I(p=0)] \cos(2\pi\tilde{\nu}p) dp.$$

Pulsed *Fourier transform* spectrometer

A pulsed *Fourier transform* spectrometer does not employ transmittance techniques. In the most general description of pulsed FT spectrometry, a sample is exposed to an energizing event which causes a periodic response. The frequency of the periodic response, as governed by the field conditions in the spectrometer, is indicative of the measured properties of the analyte.

Examples of Pulsed Fourier transform spectrometry

In magnetic spectroscopy (EPR, NMR), an RF pulse in a strong ambient magnetic field is used as the energizing event. This turns the magnetic particles at an angle to the ambient field, resulting in gyration. The gyrating spins then induce a periodic current in a detector coil. Each spin exhibits a characteristic frequency of gyration (relative to the field strength) which reveals information about the analyte.

In FT-mass spectrometry, the energizing event is the injection of the charged sample into the strong electromagnetic field of a cyclotron. These particles travel in circles, inducing a current in a fixed coil on one point in their circle. Each traveling particle exhibits a characteristic cyclotron frequency-field ratio revealing the masses in the sample.

The Free Induction Decay

Pulsed FT spectrometry gives the advantage of requiring a single, time-dependent measurement which can easily deconvolute a set of similar but distinct signals. The resulting composite signal, is called a *free induction decay*, because typically the signal will decay due to inhomogeneities in sample frequency, or simply unrecoverable loss of signal due to entropic loss of the property being measured.

Fellgett Advantage

One of the most important advantages of Fourier transform spectroscopy was shown by P.B. Fellgett, an early advocate of the method. The Fellgett advantage, also known as the multiplex principle, states that a multiplex spectrometer such as the Fourier transform spectroscopy will produce a gain of the order of the square root of m in the signal-to-noise ratio of the resulting spectrum, when compared with an equivalent scanning monochromator, where m is the number of elements comprising the resulting spectrum when the measurement noise is dominated by detector noise.

Converting spectra from time domain to frequency domain

$$S(t) = \int_{-\infty}^{\infty} I(\nu) e^{-i\nu 2\pi t} d\nu$$

The sum is performed over all contributing frequencies to give a signal $S(t)$ in the time domain.

$$I(\nu) = \int_{-\infty}^{\infty} S(t) e^{i\nu 2\pi t} dt$$

gives non-zero value when $S(t)$ contains a component that matches the oscillating function.

Remember that

$$e^{ix} = \cos x + i \sin x$$

See also

- Applied spectroscopy
- 2D-FT NMRI and Spectroscopy
- Forensic chemistry
- Forensic polymer engineering
- nuclear magnetic resonance
- Infra-red spectroscopy

Further reading

- Ellis, D.I. and Goodacre, R. (2006). "Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy". *The Analyst* **131**: 875–885. doi:10.1039/b602376m.

External links

- Description of how a Fourier transform spectrometer works ^[3]
- The Michelson or Fourier transform spectrograph ^[4]
- Internet Journal of Vibrational Spectroscopy - How FTIR works ^[5]
- Fourier Transform Spectroscopy Topical Meeting and Tabletop Exhibit ^[6]

References

- [1] Antoine Abragam. 1968. *Principles of Nuclear Magnetic Resonance.*, 895 pp., Cambridge University Press: Cambridge, UK.
- [2] Peter Atkins, Julio De Paula. 2006. *Physical Chemistry.*, 8th ed. Oxford University Press: Oxford, UK.
- [3] <http://scienceworld.wolfram.com/physics/FourierTransformSpectrometer.html>
- [4] <http://www.astro.livjm.ac.uk/courses/phys362/notes/>
- [5] <http://www.ijvs.com/volume5/edition5/section1.html#Feature>
- [6] <http://www.osa.org/meetings/topicalmeetings/fts/default.aspx>

Spectroscopy Theory

Quantum mechanics

Quantum mechanics (QM) is a set of scientific principles describing the known behavior of energy and matter that predominate at the atomic and subatomic scales. The name derives from the observation that some physical quantities—such as the energy of an electron—can be changed only by set amounts, or quanta, rather than being capable of varying by any amount. The wave–particle duality of energy and matter at the atomic scale provides a unified view of the behavior of particles such as photons and electrons. Photons are the quanta of light, and have energy values proportional to their frequency via the Planck constant. An electron bound in an atomic orbital has quantized values of angular momentum and energy. The unbound electron does not exhibit quantized energy levels, but is associated with a quantum mechanical wavelength, as are all massive particles. The full significance of the Planck constant is expressed in physics through the abstract mathematical notion of action.

The mathematical formulation of quantum mechanics is abstract and its implications are often non-intuitive. The centerpiece of this mathematical system is the wavefunction. The

wavefunction is a mathematical function of time and space that can provide information about the position and momentum of a particle, but only as probabilities, as dictated by the constraints imposed by the uncertainty principle. Mathematical manipulations of the wavefunction usually involve the bra-ket notation, which requires an understanding of complex numbers and linear functionals. Many of the results of QM can only be expressed mathematically and do not have models that are as easy to visualize as those of classical mechanics. For instance, the ground state in quantum mechanical model is a non-zero energy state that is the lowest permitted energy state of a system, rather than a more traditional system that is thought of as simple being at rest with zero kinetic energy.

Overview

The word *quantum* derives from Latin meaning "how great" or "how much".^[1] In quantum mechanics, it refers to a discrete unit that quantum theory assigns to certain physical quantities, such as the energy of an atom at rest (see Figure 1). The discovery that particles are discrete packets of energy with wave-like properties led to the branch of

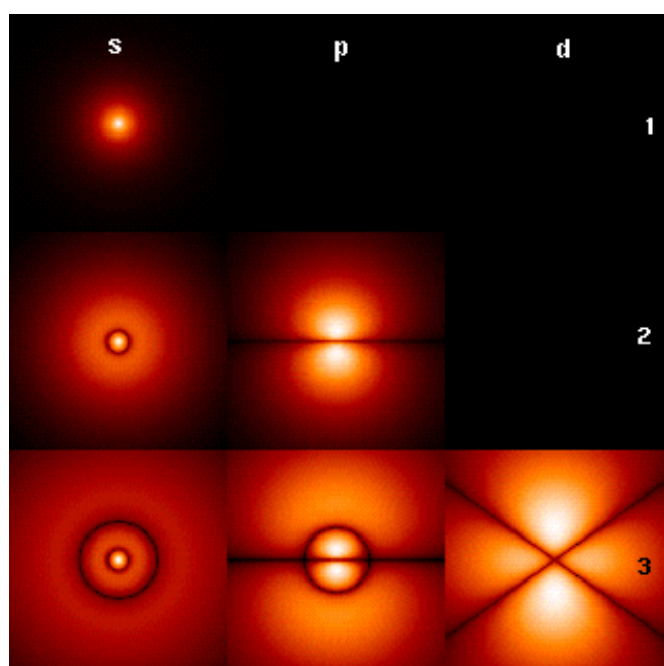


Fig. 1: Probability densities corresponding to the wavefunctions of an electron in a hydrogen atom possessing definite energy levels (increasing from the top of the image to the bottom: $n = 1, 2, 3, \dots$) and angular momentum (increasing across from left to right: $s, p, d, \dots$). Brighter areas correspond to higher probability density in a position measurement. Wavefunctions like these are directly comparable to Chladni's figures of acoustic modes of vibration classical physics and are indeed modes of oscillation as well: they possess a sharp energy and thus a keen frequency. The angular momentum and energy are quantized, and only take on discrete values like those shown (as is the case for resonant frequencies in acoustics).

physics that deals with atomic and subatomic systems which is today called quantum mechanics. It is the underlying mathematical framework of many fields of physics and chemistry, including condensed matter physics, solid-state physics, atomic physics, molecular physics, computational physics, computational chemistry, quantum chemistry, particle physics, nuclear chemistry, and nuclear physics. The foundations of quantum mechanics were established during the first half of the twentieth century by Werner Heisenberg, Max Planck, Louis de Broglie, Albert Einstein, Niels Bohr, Erwin Schrödinger, Max Born, John von Neumann, Paul Dirac, Wolfgang Pauli, David Hilbert, and others.^[2] Some fundamental aspects of the theory are still actively studied.^[3]

Quantum mechanics is essential to understand the behavior of systems at atomic length scales and smaller. For example, if classical mechanics governed the workings of an atom, electrons would rapidly travel towards and collide with the nucleus, making stable atoms impossible. However, in the natural world the electrons normally remain in an uncertain, non-deterministic "smeared" (wave–particle wave function) orbital path around or through the nucleus, defying classical electromagnetism.^[4]

Quantum mechanics was initially developed to provide a better explanation of the atom, especially the spectra of light emitted by different atomic species. The quantum theory of the atom was developed as an explanation for the electron's staying in its orbital, which could not be explained by Newton's laws of motion and by Maxwell's laws of classical electromagnetism.

In the formalism of quantum mechanics, the state of a system at a given time is described by a complex wave function (sometimes referred to as orbitals in the case of atomic electrons), and more generally, elements of a complex vector space.^[5] This abstract mathematical object allows for the calculation of probabilities of outcomes of concrete experiments. For example, it allows one to compute the probability of finding an electron in a particular region around the nucleus at a particular time. Contrary to classical mechanics, one can never make simultaneous predictions of conjugate variables, such as position and momentum, with accuracy. For instance, electrons may be considered to be located somewhere within a region of space, but with their exact positions being unknown. Contours of constant probability, often referred to as "clouds", may be drawn around the nucleus of an atom to conceptualize where the electron might be located with the most probability. Heisenberg's uncertainty principle quantifies the inability to precisely locate the particle given its conjugate.^[6]

The other exemplar that led to quantum mechanics was the study of electromagnetic waves such as light. When it was found in 1900 by Max Planck that the energy of waves could be described as consisting of small packets or quanta, Albert Einstein further developed this idea to show that an electromagnetic wave such as light could be described by a particle called the photon with a discrete energy dependent on its frequency. This led to a theory of unity between subatomic particles and electromagnetic waves called wave–particle duality in which particles and waves were neither one nor the other, but had certain properties of both. While quantum mechanics describes the world of the very small, it also is needed to explain certain macroscopic quantum systems such as superconductors and superfluids.

Broadly speaking, quantum mechanics incorporates four classes of phenomena for which classical physics cannot account: (I) the quantization (discretization) of certain physical quantities, (II) wave–particle duality, (III) the uncertainty principle, and (IV) quantum entanglement. Each of these phenomena is described in detail in subsequent sections.

History

The history of quantum mechanics began with the 1838 discovery of cathode rays by Michael Faraday, the 1859 statement of the black body radiation problem by Gustav Kirchhoff, the 1877 suggestion by Ludwig Boltzmann that the energy states of a physical system could be discrete, and the 1900 quantum hypothesis by Max Planck.^[7] Planck's hypothesis stated that any energy is radiated and absorbed in quantities divisible by discrete "energy elements", such that each energy element E is proportional to its frequency ν :

$$E = h\nu$$

where h is Planck's action constant. Planck insisted that this was simply an aspect of the processes of absorption and emission of radiation and had nothing to do with the physical reality of the radiation itself.^[8] However, at that time, this appeared not to explain the photoelectric effect (1839), i.e. that shining light on certain materials can eject electrons from the material. In 1905, basing his work on Planck's quantum hypothesis, Albert Einstein postulated that light itself consists of individual quanta.^[9]

In the mid-1920s, developments in quantum mechanics quickly led to it becoming the standard formulation for atomic physics. In the summer of 1925, Bohr and Heisenberg published results that closed the "Old Quantum Theory". Light quanta came to be called photons (1926). From Einstein's simple postulation was born a flurry of debating, theorizing and testing, and thus, the entire field of quantum physics, leading to its wider acceptance at the Fifth Solvay Conference in 1927.

Quantum mechanics and classical physics

Predictions of quantum mechanics have been verified experimentally to a very high degree of accuracy. Thus, the current logic of correspondence principle between classical and quantum mechanics is that all objects obey laws of quantum mechanics, and classical mechanics is just a quantum mechanics of large systems (or a statistical quantum mechanics of a large collection of particles). Laws of classical mechanics thus follow from laws of quantum mechanics at the limit of large systems or large quantum numbers.^[10] However, chaotic systems do not have good quantum numbers, and quantum chaos studies the relationship between classical and quantum descriptions in these systems.

The main differences between classical and quantum theories have already been mentioned above in the remarks on the Einstein-Podolsky-Rosen paradox. Essentially the difference boils down to the statement that quantum mechanics is coherent (addition of *amplitudes*), whereas classical theories are incoherent (addition of *intensities*). Thus, such quantities as *coherence lengths* and *coherence times* come into play. For microscopic bodies the extension of the system is certainly much smaller than the coherence length; for macroscopic bodies one expects that it should be the other way round.^[11] An exception to this rule can occur at extremely low temperatures, when quantum behavior can manifest itself on more macroscopic scales (see Bose-Einstein condensate).

This is in accordance with the following observations:

Many macroscopic properties of classical systems are direct consequences of quantum behavior of its parts. For example, the stability of bulk matter (which consists of atoms and molecules which would quickly collapse under electric forces alone), the rigidity of solids, and the mechanical, thermal, chemical, optical and magnetic properties of matter are all results of interaction of electric charges under the rules of quantum mechanics.^[12]

While the seemingly exotic behavior of matter posited by quantum mechanics and relativity theory become more apparent when dealing with extremely fast-moving or extremely tiny particles, the laws of classical Newtonian physics remain accurate in predicting the behavior of large objects—of the order of the size of large molecules and bigger—at velocities much smaller than the velocity of light.^[13]

Theory

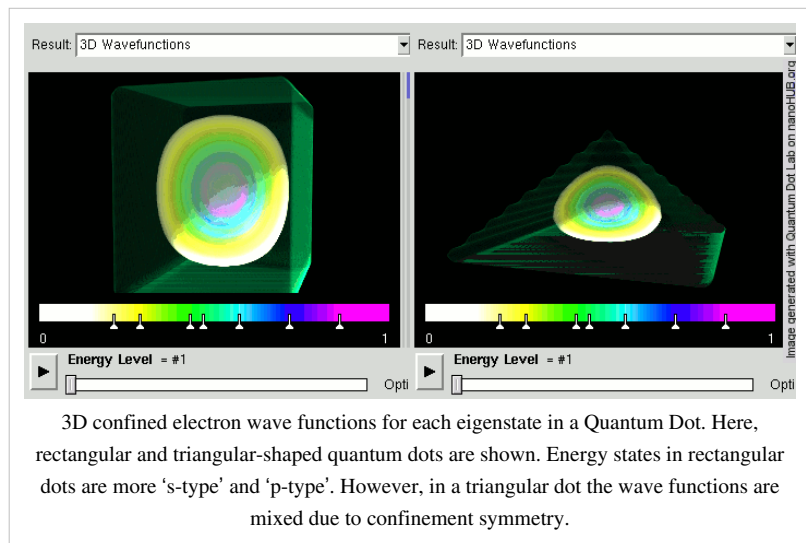
There are numerous mathematically equivalent formulations of quantum mechanics. One of the oldest and most commonly used formulations is the transformation theory proposed by Cambridge theoretical physicist Paul Dirac, which unifies and generalizes the two earliest formulations of quantum mechanics, matrix mechanics (invented by Werner Heisenberg)^{[14] [15]} and wave mechanics (invented by Erwin Schrödinger).^[16]

In this formulation, the instantaneous state of a quantum system encodes the probabilities of its measurable properties, or "observables". Examples of observables include energy, position, momentum, and angular momentum. Observables can be either continuous (e.g., the position of a particle) or discrete (e.g., the energy of an electron bound to a hydrogen atom).^[17] Generally, quantum mechanics does not assign definite values to observables. Instead, it makes predictions using probability distributions; that is, the probability of obtaining possible outcomes from measuring an observable. Oftentimes these results are skewed by many causes, such as dense probability clouds^[18] or quantum state nuclear attraction.^{[19] [20]} Naturally, these probabilities will depend on the quantum state at the "instant" of the measurement. Hence, uncertainty is involved in the value. There are, however, certain states that are associated with a definite value of a particular observable. These are known as eigenstates of the observable ("eigen" can be translated from German as inherent or as a characteristic).^[21] In the everyday world, it is natural and intuitive to think of everything (every observable) as being in an eigenstate. Everything appears to have a definite position, a definite momentum, a definite energy, and a definite time of occurrence. However, quantum mechanics does not pinpoint the exact values of a particle for its position and momentum (since they are conjugate pairs) or its energy and time (since they too are conjugate pairs); rather, it only provides a range of probabilities of where that particle might be given its momentum and momentum probability. Therefore, it is helpful to use different words to describe states having *uncertain* values and states having *definite* values (eigenstate).

For example, consider a free particle. In quantum mechanics, there is wave-particle duality so the properties of the particle can be described as the properties of a wave. Therefore, its quantum state can be represented as a wave of arbitrary shape and extending over space as a wave function. The position and momentum of the particle are observables. The Uncertainty Principle states that both the position and the momentum cannot simultaneously be measured with full precision at the same time. However, one can measure the position alone of a

moving free particle creating an eigenstate of position with a wavefunction that is very large (a Dirac delta) at a particular position x and zero everywhere else. If one performs a position measurement on such a wavefunction, the result x will be obtained with 100% probability (full certainty). This is called an eigenstate of position (mathematically more precise: a *generalized position eigenstate (eigendistribution)*). If the particle is in an eigenstate of position then its momentum is completely unknown. On the other hand, if the particle is in an eigenstate of momentum then its position is completely unknown.^[22] In an eigenstate of momentum having a plane wave form, it can be shown that the wavelength is equal to h/p , where h is Planck's constant and p is the momentum of the eigenstate.^[23]

Usually, a system will not be in an eigenstate of the observable we are interested in. However, if one measures the observable, the wavefunction will instantaneously be an eigenstate (or generalized eigenstate) of that observable.



This process is known as wavefunction collapse, a debatable process.^[24] It involves expanding the system under study to include the measurement device. If one knows the corresponding wave function at the instant before the measurement, one will be able to compute the probability of collapsing into each of the possible eigenstates. For example, the free particle in the previous example will usually have a wavefunction that is a wave packet centered around some mean position x_0 , neither an eigenstate of position nor of momentum. When one measures the position of the particle, it is impossible to predict with certainty the result.^[25] It is probable, but not certain, that it will be near x_0 , where the amplitude of the wave function is large. After the measurement is performed, having obtained some result x , the wave function collapses into a position eigenstate centered at x .^[26]

Wave functions can change as time progresses. An equation known as the Schrödinger equation describes how wave functions change in time, a role similar to Newton's second law in classical mechanics. The Schrödinger equation, applied to the aforementioned example of the free particle, predicts that the center of a wave packet will move through space at a constant velocity, like a classical particle with no forces acting on it. However, the wave packet will also spread out as time progresses, which means that the position becomes more uncertain. This also has the effect of turning position eigenstates (which can be thought of as infinitely sharp wave packets) into broadened wave packets that are no longer position eigenstates.^[27] Some wave functions produce probability distributions that are constant or independent of time, such as when in a stationary state of constant energy, time drops out of the absolute square of the wave function. Many systems that are treated dynamically in classical mechanics are described by such "static" wave functions. For example, a single electron in an unexcited atom is pictured classically as a particle moving in a circular trajectory around the atomic nucleus, whereas in quantum mechanics it is described by a static, spherically symmetric wavefunction surrounding the nucleus (Fig. 1). (Note that only the lowest angular momentum states, labeled s , are spherically symmetric).^[28]

The time evolution of wave functions is deterministic in the sense that, given a wavefunction at an initial time, it makes a definite prediction of what the wavefunction will be at any later time.^[29] During a measurement, the change of the wavefunction into another one is not deterministic, but rather unpredictable, i.e., random. A time-evolution simulation can be seen here.^[30]

The probabilistic nature of quantum mechanics thus stems from the act of measurement. This is one of the most difficult aspects of quantum systems to understand. It was the central topic in the famous Bohr-Einstein debates, in which the two scientists attempted to clarify these fundamental principles by way of thought experiments. In the decades after the formulation of quantum mechanics, the question of what constitutes a "measurement" has been extensively studied. Interpretations of quantum mechanics have been formulated to do away with the concept of "wavefunction collapse"; see, for example, the relative state interpretation. The basic idea is that when a quantum system interacts with a measuring apparatus, their respective wavefunctions become entangled, so that the original quantum system ceases to exist as an independent entity. For details, see the article on measurement in quantum mechanics.^[31]

Mathematical formulation

In the mathematically rigorous formulation of quantum mechanics, developed by Paul Dirac^[32] and John von Neumann,^[33] the possible states of a quantum mechanical system are represented by unit vectors (called "state vectors") residing in a complex separable Hilbert space (variously called the "state space" or the "associated Hilbert space" of the system) well defined up to a complex number of norm 1 (the phase factor). In other words, the possible states are points in the projectivization of a Hilbert space, usually called the complex projective space. The exact nature of this Hilbert space is dependent on the system; for example, the state space for position and momentum states is the space of square-integrable functions, while the state space for the spin of a single proton is just the product of two complex planes. Each observable is represented by a maximally-Hermitian (precisely: by a self-adjoint) linear operator acting on the state space. Each eigenstate of an observable corresponds to an eigenvector of the operator, and the associated eigenvalue corresponds to the value of the observable in that eigenstate. If the

operator's spectrum is discrete, the observable can only attain those discrete eigenvalues.

The time evolution of a quantum state is described by the Schrödinger equation, in which the Hamiltonian, the operator corresponding to the total energy of the system, generates time evolution.

The inner product between two state vectors is a complex number known as a probability amplitude. During a measurement, the probability that a system collapses from a given initial state to a particular eigenstate is given by the square of the absolute value of the probability amplitudes between the initial and final states. The possible results of a measurement are the eigenvalues of the operator — which explains the choice of *Hermitian* operators, for which all the eigenvalues are real. We can find the probability distribution of an observable in a given state by computing the spectral decomposition of the corresponding operator. Heisenberg's uncertainty principle is represented by the statement that the operators corresponding to certain observables do not commute.

The Schrödinger equation acts on the entire probability amplitude, not merely its absolute value. Whereas the absolute value of the probability amplitude encodes information about probabilities, its phase encodes information about the interference between quantum states. This gives rise to the wave-like behavior of quantum states.

It turns out that analytic solutions of Schrödinger's equation are only available for a small number of model Hamiltonians, of which the quantum harmonic oscillator, the particle in a box, the hydrogen molecular ion and the hydrogen atom are the most important representatives. Even the helium atom, which contains just one more electron than hydrogen, defies all attempts at a fully analytic treatment. There exist several techniques for generating approximate solutions. For instance, in the method known as perturbation theory one uses the analytic results for a simple quantum mechanical model to generate results for a more complicated model related to the simple model by, for example, the addition of a weak potential energy. Another method is the "semi-classical equation of motion" approach, which applies to systems for which quantum mechanics produces weak deviations from classical behavior. The deviations can be calculated based on the classical motion. This approach is important for the field of quantum chaos.

An alternative formulation of quantum mechanics is Feynman's path integral formulation, in which a quantum-mechanical amplitude is considered as a sum over histories between initial and final states; this is the quantum-mechanical counterpart of action principles in classical mechanics.

Interactions with other scientific theories

The fundamental rules of quantum mechanics are very deep. They assert that the state space of a system is a Hilbert space and the observables are Hermitian operators acting on that space, but do not tell us which Hilbert space or which operators, or if it even exists. These must be chosen appropriately in order to obtain a quantitative description of a quantum system. An important guide for making these choices is the correspondence principle, which states that the predictions of quantum mechanics reduce to those of classical physics when a system moves to higher energies or equivalently, larger quantum numbers. In other words, classical mechanics is simply a quantum mechanics of large systems. This "high energy" limit is known as the *classical* or *correspondence limit*. One can therefore start from an established classical model of a particular system, and attempt to guess the underlying quantum model that gives rise to the classical model in the correspondence limit.

Unsolved problems in physics

*In the correspondence limit of **quantum mechanics**: Is there a preferred interpretation of quantum mechanics? How does the quantum description of reality, which includes elements such as the "superposition of states" and "wavefunction collapse", give rise to the reality we perceive?*



When quantum mechanics was originally formulated, it was applied to models whose correspondence limit was non-relativistic classical mechanics. For instance, the well-known model of the quantum harmonic oscillator uses an explicitly non-relativistic expression for the kinetic energy of the oscillator, and is thus a quantum version of the classical harmonic oscillator.

Early attempts to merge quantum mechanics with special relativity involved the replacement of the Schrödinger equation with a covariant equation such as the Klein-Gordon equation or the Dirac equation. While these theories were successful in explaining many experimental results, they had certain unsatisfactory qualities stemming from their neglect of the relativistic creation and annihilation of particles. A fully relativistic quantum theory required the development of quantum field theory, which applies quantization to a field rather than a fixed set of particles. The first complete quantum field theory, quantum electrodynamics, provides a fully quantum description of the electromagnetic interaction.

The full apparatus of quantum field theory is often unnecessary for describing electrodynamic systems. A simpler approach, one employed since the inception of quantum mechanics, is to treat charged particles as quantum mechanical objects being acted on by a classical electromagnetic field. For example, the elementary quantum model of the hydrogen atom describes the electric field of the hydrogen atom using a classical $-\frac{e^2}{4\pi\epsilon_0} \frac{1}{r}$ Coulomb potential. This "semi-classical" approach fails if quantum fluctuations in the electromagnetic field play an important role, such as in the emission of photons by charged particles.

Quantum field theories for the strong nuclear force and the weak nuclear force have been developed. The quantum field theory of the strong nuclear force is called quantum chromodynamics, and describes the interactions of the subnuclear particles: quarks and gluons. The weak nuclear force and the electromagnetic force were unified, in their quantized forms, into a single quantum field theory known as electroweak theory, by the physicists Abdus Salam, Sheldon Glashow and Steven Weinberg. These three men shared the Nobel Prize in Physics in 1979 for this work.^[34]

It has proven difficult to construct quantum models of gravity, the remaining fundamental force. Semi-classical approximations are workable, and have led to predictions such as Hawking radiation. However, the formulation of a complete theory of quantum gravity is hindered by apparent incompatibilities between general relativity, the most accurate theory of gravity currently known, and some of the fundamental assumptions of quantum theory. The resolution of these incompatibilities is an area of active research, and theories such as string theory are among the possible candidates for a future theory of quantum gravity.

In the 21st century classical mechanics has been extended into the complex domain and complex classical mechanics exhibits behaviours very similar to quantum mechanics.^[35]

Example

The particle in a 1-dimensional potential energy box is the most simple example where restraints lead to the quantization of energy levels. The box is defined as zero potential energy inside a certain interval and infinite everywhere outside that interval. For the 1-dimensional case in the x direction, the time-independent Schrödinger equation can be written as:^[36]

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} = E\psi.$$

The general solutions are:

$$\psi(x) = Ae^{ikx} + Be^{-ikx} \quad E = \frac{\hbar^2 k^2}{2m}$$

or, from Euler's formula,

$$\psi(x) = C \sin kx + D \cos kx.$$

The presence of the walls of the box determines the values of C , D , and k . At each wall ($x = 0$ and $x = L$), $\psi = 0$. Thus when $x = 0$,

$$\psi(0) = 0 = C \sin 0 + D \cos 0 = D$$

and so $D = 0$. When $x = L$,

$$\psi(L) = 0 = C \sin kL.$$

C cannot be zero, since this would conflict with the Born interpretation. Therefore $\sin kL = 0$, and so it must be that kL is an integer multiple of π . Therefore,

$$k = \frac{n\pi}{L} \quad n = 1, 2, 3, \dots$$

The quantization of energy levels follows from this constraint on k , since

$$E = \frac{\hbar^2 \pi^2 n^2}{2mL^2} = \frac{n^2 h^2}{8mL^2}.$$

Attempts at a unified field theory

As of 2010 the quest for unifying the fundamental forces through quantum mechanics is still ongoing. Quantum electrodynamics (or "quantum electromagnetism"), which is currently the most accurately tested physical theory,^[37] has been successfully merged with the weak nuclear force into the electroweak force and work is currently being done to merge the electroweak and strong force into the electrostrong force. Current predictions state that at around 10^{14} GeV the three aforementioned forces are fused into a single unified field,^[38] Beyond this "grand unification", it is speculated that it may be possible to merge gravity with the other three gauge symmetries, expected to occur at roughly 10^{19} GeV. However — and while special relativity is parsimoniously incorporated into quantum electrodynamics — the expanded general relativity, currently the best theory describing the gravitation force, has not been fully incorporated into quantum theory.

Relativity and quantum mechanics

Main articles: Quantum gravity and Theory of everything

Even with the defining postulates of both Einstein's theory of general relativity and quantum theory being indisputably supported by rigorous and repeated empirical evidence and while they do not directly contradict each other theoretically (at least with regard to primary claims), they are resistant to being incorporated within one cohesive model.^[39]

Einstein himself is well known for rejecting some of the claims of quantum mechanics. While clearly contributing to the field, he did not accept the more philosophical consequences and interpretations of quantum mechanics, such as the lack of deterministic causality and the assertion that a single subatomic particle can occupy numerous areas of space at one time. He also was the first to notice some of the apparently exotic consequences of entanglement and used them to formulate the Einstein-Podolsky-Rosen paradox, in the hope of showing that quantum mechanics had unacceptable implications. This was 1935, but in 1964 it was shown by John Bell (see Bell inequality) that Einstein's assumption was correct, but had to be completed by *hidden variables* and thus based on wrong philosophical assumptions. According to the paper of J. Bell and the Copenhagen interpretation (the common interpretation of quantum mechanics by physicists since 1927), and contrary to Einstein's ideas, quantum mechanics was

- neither a "realistic" theory (since quantum measurements do not *state* pre-existing properties, but rather they *prepare* properties)
- nor a *local* theory (essentially not, because the state vector $|\psi\rangle$ determines simultaneously the probability amplitudes at all sites, $|\psi\rangle \rightarrow \psi(\mathbf{r}), \forall \mathbf{r}$).

The Einstein-Podolsky-Rosen paradox shows in any case that there exist experiments by which one can measure the state of one particle and instantaneously change the state of its entangled partner, although the two particles can be an arbitrary distance apart; however, this effect does not violate causality, since no transfer of information happens. These experiments are the basis of some of the most topical applications of the theory, quantum cryptography, which has been on the market since 2004 and works well, although at small distances of typically ≤ 1000 km.

Gravity is negligible in many areas of particle physics, so that unification between general relativity and quantum mechanics is not an urgent issue in those applications. However, the lack of a correct theory of quantum gravity is an important issue in cosmology and physicists' search for an elegant "theory of everything". Thus, resolving the inconsistencies between both theories has been a major goal of twentieth- and twenty-first-century physics. Many prominent physicists, including Stephen Hawking, have labored in the attempt to discover a theory underlying *everything*, combining not only different models of subatomic physics, but also deriving the universe's four forces—the strong force, electromagnetism, weak force, and gravity—from a single force or phenomenon. One of the leaders in this field is Edward Witten, a theoretical physicist who formulated the groundbreaking M-theory, which is an attempt at describing the supersymmetrical based string theory.

Applications

Quantum mechanics has had enormous success in explaining many of the features of our world. The individual behaviour of the subatomic particles that make up all forms of matter—electrons, protons, neutrons, photons and others—can often only be satisfactorily described using quantum mechanics. Quantum mechanics has strongly influenced string theory, a candidate for a theory of everything (see reductionism) and the multiverse hypothesis. It is also related to statistical mechanics.

Quantum mechanics is important for understanding how individual atoms combine covalently to form chemicals or molecules. The application of quantum mechanics to chemistry is known as quantum chemistry. (Relativistic) quantum mechanics can in principle mathematically describe most of chemistry. Quantum mechanics can provide quantitative insight into ionic and covalent bonding processes by explicitly showing which molecules are energetically favorable to which others, and by approximately how much.^[40] Most of the calculations performed in

computational chemistry rely on quantum mechanics.^[41]

Much of modern technology operates at a scale where quantum effects are significant. Examples include the laser, the transistor (and thus the microchip), the electron microscope, and magnetic resonance imaging. The study of semiconductors led to the invention of the diode and the transistor, which are indispensable for modern electronics.

Researchers are currently seeking robust methods of directly manipulating quantum states. Efforts are being made to develop quantum cryptography, which will allow guaranteed secure transmission of information. A more distant goal is the development of quantum computers, which are expected to perform certain computational tasks exponentially faster than classical computers. Another active research topic is quantum teleportation, which deals with techniques to transmit quantum states over arbitrary distances.

Quantum tunneling is vital in many devices, even in the simple light switch, as otherwise the electrons in the electric current could not penetrate the potential barrier made up of a layer of oxide. Flash memory chips found in USB drives use quantum tunneling to erase their memory cells.

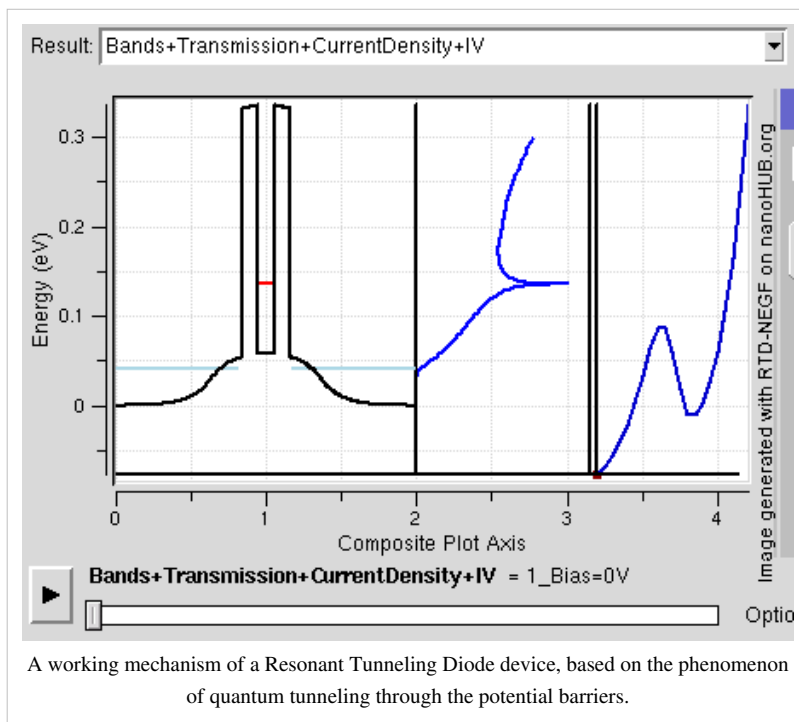
QM primarily applies to the atomic regimes of matter and energy, but some systems exhibit quantum mechanical effects on a large scale; superfluidity (the frictionless flow of a liquid at temperatures near absolute zero) is one well-known example. Quantum theory also provides accurate descriptions for many previously unexplained phenomena such as black body radiation and the stability of electron orbitals. It has also given insight into the workings of many different biological systems, including smell receptors and protein structures.^[42] Even so, classical physics often can be a good approximation to results otherwise obtained by **quantum physics**, typically in circumstances with large numbers of particles or large quantum numbers. (However, some open questions remain in the field of quantum chaos.)

Philosophical consequences

Since its inception, the many counter-intuitive results of quantum mechanics have provoked strong philosophical debate and many interpretations. Even fundamental issues such as Max Born's basic rules concerning probability amplitudes and probability distributions took decades to be appreciated.

The Copenhagen interpretation, due largely to the Danish theoretical physicist Niels Bohr, is the interpretation of quantum mechanics most widely accepted amongst physicists. According to it, the probabilistic nature of quantum mechanics predictions cannot be explained in terms of some other deterministic theory, and does not simply reflect our limited knowledge. Quantum mechanics provides probabilistic results because the physical universe is itself probabilistic rather than deterministic.

Albert Einstein, himself one of the founders of quantum theory, disliked this loss of determinism in measurement (this dislike is the source of his famous quote, "God does not play dice with the universe."). Einstein held that there



should be a local hidden variable theory underlying quantum mechanics and that, consequently, the present theory was incomplete. He produced a series of objections to the theory, the most famous of which has become known as the Einstein-Podolsky-Rosen paradox. John Bell showed that the EPR paradox led to experimentally testable differences between quantum mechanics and local realistic theories. Experiments have been performed confirming the accuracy of quantum mechanics, thus demonstrating that the physical world cannot be described by local realistic theories.^[43] The *Bohr-Einstein debates* provide a vibrant critique of the Copenhagen Interpretation from an epistemological point of view.

The Everett many-worlds interpretation, formulated in 1956, holds that all the possibilities described by quantum theory simultaneously occur in a multiverse composed of mostly independent parallel universes.^[44] This is not accomplished by introducing some new axiom to quantum mechanics, but on the contrary by *removing* the axiom of the collapse of the wave packet: All the possible consistent states of the measured system and the measuring apparatus (including the observer) are present in a *real* physical (not just formally mathematical, as in other interpretations) quantum superposition. Such a superposition of consistent state combinations of different systems is called an entangled state.

While the multiverse is deterministic, we perceive non-deterministic behavior governed by probabilities, because we can observe only the universe, i.e. the consistent state contribution to the mentioned superposition, we inhabit. Everett's interpretation is perfectly consistent with John Bell's experiments and makes them intuitively understandable. However, according to the theory of quantum decoherence, the parallel universes will never be accessible to us. This inaccessibility can be understood as follows: Once a measurement is done, the measured system becomes entangled with both the physicist who measured it and a huge number of other particles, some of which are photons flying away towards the other end of the universe; in order to prove that the wave function did not collapse one would have to bring all these particles back and measure them again, together with the system that was measured originally. This is completely impractical, but even if one could theoretically do this, it would destroy any evidence that the original measurement took place (including the physicist's memory).

See also

- Copenhagen interpretation
 - Correspondence rules
 - De Broglie–Bohm theory
 - EPR paradox
 - Fine-structure constant
 - Interpretation of quantum mechanics
 - Introduction to quantum mechanics
 - Many-worlds interpretation
 - Measurement in quantum mechanics
 - Measurement problem
 - Photon dynamics in the double-slit experiment
 - Photon polarization
 - Physical ontology
 - Quantum chaos
 - Quantum chemistry
 - Quantum chemistry computer programs
 - Quantum chromodynamics
 - Quantum computers
 - Quantum decoherence
 - Quantum electrochemistry
-

- Quantum electronics
- Quantum field theory
- Quantum information
- Quantum mind^[45]
- Quantum optics
- Quantum pseudo-telepathy
- Quantum thermodynamics
- Quantum triviality
- Quantum Zeno effect
- Quasi-set theory
- Relation between Schrödinger's equation and the path integral formulation of quantum mechanics
- Schrödinger's cat
- Theoretical and experimental justification for the Schrödinger equation
- Theoretical chemistry
- Transactional interpretation
- Trojan wave packet

References

The following titles, all by working physicists, attempt to communicate quantum theory to lay people, using a minimum of technical apparatus.

- Chester, Marvin (1987) *Primer of Quantum Mechanics*. John Wiley. ISBN 0-486-42878-8
- Richard Feynman, 1985. *QED: The Strange Theory of Light and Matter*, Princeton University Press. ISBN 0-691-08388-6. Four elementary lectures on quantum electrodynamics and quantum field theory, yet containing many insights for the expert.
- Ghirardi, GianCarlo, 2004. *Sneaking a Look at God's Cards*, Gerald Malsbary, trans. Princeton Univ. Press. The most technical of the works cited here. Passages using algebra, trigonometry, and bra-ket notation can be passed over on a first reading.
- N. David Mermin, 1990, "Spooky actions at a distance: mysteries of the QT" in his *Boojums all the way through*. Cambridge University Press: 110-76.
- Victor Stenger, 2000. *Timeless Reality: Symmetry, Simplicity, and Multiple Universes*. Buffalo NY: Prometheus Books. Chpts. 5-8. Includes cosmological and philosophical considerations.

More technical:

- Bryce DeWitt, R. Neill Graham, eds., 1973. *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton Series in Physics, Princeton University Press. ISBN 0-691-08131-X
- Dirac, P. A. M. (1930). *The Principles of Quantum Mechanics*. ISBN 0198520115. The beginning chapters make up a very clear and comprehensible introduction.
- Hugh Everett, 1957, "Relative State Formulation of Quantum Mechanics," *Reviews of Modern Physics* 29: 454-62.
- Feynman, Richard P.; Leighton, Robert B.; Sands, Matthew (1965). *The Feynman Lectures on Physics*. **1-3**. Addison-Wesley. ISBN 0738200085.
- Griffiths, David J. (2004). *Introduction to Quantum Mechanics (2nd ed.)*. Prentice Hall. ISBN 0-13-111892-7. OCLC 40251748. A standard undergraduate text.
- Max Jammer, 1966. *The Conceptual Development of Quantum Mechanics*. McGraw Hill.
- Hagen Kleinert, 2004. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*, 3rd ed. Singapore: World Scientific. Draft of 4th edition.^[46]
- Gunther Ludwig, 1968. *Wave Mechanics*. London: Pergamon Press. ISBN 0-08-203204-1

- George Mackey (2004). *The mathematical foundations of quantum mechanics*. Dover Publications. ISBN 0-486-43517-2.
- Albert Messiah, 1966. *Quantum Mechanics* (Vol. I), English translation from French by G. M. Temmer. North Holland, John Wiley & Sons. Cf. chpt. IV, section III.
- Omnès, Roland (1999). *Understanding Quantum Mechanics*. Princeton University Press. ISBN 0-691-00435-8. OCLC 39849482.
- Scerri, Eric R., 2006. *The Periodic Table: Its Story and Its Significance*. Oxford University Press. Considers the extent to which chemistry and the periodic system have been reduced to quantum mechanics. ISBN 0-19-530573-6
- Transnational College of Lex (1996). *What is Quantum Mechanics? A Physics Adventure*. Language Research Foundation, Boston. ISBN 0-9643504-1-6. OCLC 34661512.
- von Neumann, John (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton University Press. ISBN 0691028931.
- Hermann Weyl, 1950. *The Theory of Groups and Quantum Mechanics*, Dover Publications.
- D. Greenberger, K. Hentschel, F. Weinert, eds., 2009. *Compendium of quantum physics, Concepts, experiments, history and philosophy*, Springer-Verlag, Berlin, Heidelberg.

Further reading

- Bohm, David (1989). *Quantum Theory*. Dover Publications. ISBN 0-486-65969-0.
- Eisberg, Robert; Resnick, Robert (1985). *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles* (2nd ed.). Wiley. ISBN 0-471-87373-X.
- Liboff, Richard L. (2002). *Introductory Quantum Mechanics*. Addison-Wesley. ISBN 0-8053-8714-5.
- Merzbacher, Eugen (1998). *Quantum Mechanics*. Wiley, John & Sons, Inc. ISBN 0-471-88702-1.
- Sakurai, J. J. (1994). *Modern Quantum Mechanics*. Addison Wesley. ISBN 0-201-53929-2.
- Shankar, R. (1994). *Principles of Quantum Mechanics*. Springer. ISBN 0-306-44790-8.

External links

General

- The Modern Revolution in Physics ^[47] - an online textbook.
- J. O'Connor and E. F. Robertson: A history of quantum mechanics. ^[48]
- Introduction to Quantum Theory at Quantiki. ^[49]
- Quantum Physics Made Relatively Simple ^[50]: three video lectures by Hans Bethe
- H is for h-bar. ^[51]
- Quantum Mechanics Books Collection ^[52]: Collection of free books

Course material

- Doron Cohen: Lecture notes in Quantum Mechanics (comprehensive, with advanced topics). ^[53]
- MIT OpenCourseWare: Chemistry ^[54]. See 5.61 ^[55], 5.73 ^[56], and 5.74 ^[57]
- MIT OpenCourseWare: Physics ^[58]. See 8.04 ^[59], 8.05 ^[60], and 8.06 ^[61]
- Stanford Continuing Education PHY 25: Quantum Mechanics ^[62] by Leonard Susskind, see course description ^[63] Fall 2007
- 5½ Examples in Quantum Mechanics ^[64]
- Imperial College Quantum Mechanics Course. ^[65]
- Spark Notes - Quantum Physics. ^[66]
- Quantum Physics Online : interactive introduction to quantum mechanics (RS applets). ^[67]
- Experiments to the foundations of quantum physics with single photons. ^[68]
- Motion Mountain, Volume IV ^[69] - A modern introduction to quantum theory, with several animations.

- AQME^[70]: Advancing Quantum Mechanics for Engineers — by T.Barzso, D.Vasileska and G.Klimeck online learning resource with simulation tools on nanoHUB
- Quantum Mechanics^[71] by Martin Plenio
- Quantum Mechanics^[72] by Richard Fitzpatrick

FAQs

- Many-worlds or relative-state interpretation.^[73]
- Measurement in Quantum mechanics.^[74]

Media

- Everything you wanted to know about the quantum world^[75] — archive of articles from *New Scientist*.
- Quantum Physics Research^[76] from Science Daily
- "Quantum Trickery: Testing Einstein's Strangest Theory"^[77]. *The New York Times*. December 27, 2005.

Philosophy

- "'Quantum Mechanics"^[78]" article by Jenann Ismael. in the *Stanford Encyclopedia of Philosophy*
- "'Measurement in Quantum Theory"^[78]" article by Henry Krips. in the *Stanford Encyclopedia of Philosophy*

References

- [1] Merriam-Webster.com (<http://www.merriam-webster.com/dictionary/quantum>)
- [2] FCCJ.org (<http://mooni.fccj.org/~ethall/quantum/quant.htm>)
- [3] Compare the list of conferences presented here (<http://ysfine.com>).
- [4] Oocities.com (http://www.oocities.com/mik_malm/quantmech.html)
- [5] Greiner, Walter; Müller, Berndt (1994). *Quantum Mechanics Symmetries, Second edition* (<http://books.google.com/books?id=gCfvWx6vuzUC&pg=PA52>). Springer-Verlag. p. 52. ISBN 3-540-58080-8. .,
- [6] AIP.org (<http://www.aip.org/history/heisenberg/p08a.htm>)
- [7] J. Mehra and H. Rechenberg, *The historical development of quantum theory*, Springer-Verlag, 1982.
- [8] T.S. Kuhn, *Black-body theory and the quantum discontinuity 1894-1912*, Clarendon Press, Oxford, 1978.
- [9] A. Einstein, *Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (On a heuristic point of view concerning the production and transformation of light)*, Annalen der Physik **17** (1905) 132-148 (reprinted in *The collected papers of Albert Einstein*, John Stachel, editor, Princeton University Press, 1989, Vol. 2, pp. 149-166, in German; see also *Einstein's early work on the quantum hypothesis*, *ibid.* pp. 134-148).
- [10] Scribd.com (<http://www.scribd.com/doc/5998949/Quantum-mechanics-course-iwhatisquantummechanics>)
- [11] Philsci-archive.pitt.edu (<http://philsci-archive.pitt.edu/archive/00002328/01/handbook.pdf>)
- [12] Academic.brooklyn.cuny.edu (<http://academic.brooklyn.cuny.edu/physics/sobel/Nucphys/atomprop.html>)
- [13] Cambridge.org (http://assets.cambridge.org/9780521829526/excerpt/9780521829526_excerpt.pdf)
- [14] Spaceandmotion.com (<http://www.spaceandmotion.com/physics-quantum-mechanics-werner-heisenberg.htm>)
- [15] Especially since Werner Heisenberg was awarded the Nobel Prize in Physics in 1932 for the creation of quantum mechanics, the role of Max Born has been obfuscated. A 2005 biography of Born details his role as the creator of the matrix formulation of quantum mechanics. This was recognized in a paper by Heisenberg, in 1940, honoring Max Planck. See: Nancy Thorndike Greenspan, "The End of the Certain World: The Life and Science of Max Born" (Basic Books, 2005), pp. 124 - 128, and 285 - 286.
- [16] IF.uj.edu.pl (<http://th-www.if.uj.edu.pl/acta/vol19/pdf/v19p0683.pdf>)
- [17] OCW.ssu.edu (http://ocw.ssu.edu/physics/classical-mechanics/pdf_lectures/06.pdf)
- [18] probability clouds are approximate, but better than the Bohr model, whereby electron location is given by a probability function, the wave function eigenvalue, such that the probability is the squared modulus of the complex amplitude
- [19] Actapress.com (<http://www.actapress.com/PaperInfo.aspx?PaperID=25988&reason=500>)
- [20] Hirshleifer, Jack (2001). *The Dark Side of the Force: Economic Foundations of Conflict Theory* (<http://books.google.com/books?id=W2J2IXgiZVgC&pg=PA265>). Cambridge University Press. p. 265. ISBN 0-521-80412-4. .,
- [21] Dict.cc (<http://www.dict.cc/german-english/eigen.html>)
De.pons.eu (<http://de.pons.eu/deutsch-englisch/eigen>)
- [22] Davies, P. C. W.; Betts, David S. (1984). *Quantum Mechanics, Second edition* (<http://books.google.com/books?id=XRyHCrGNStoC&pg=PA79>). Chapman and Hall. p. 79. ISBN 0-7487-4446-0. .,
- [23] Books.Google.com (http://books.google.com/books?id=tKm-Ekwke_UC)
- [24] PHY.olemiss.edu (<http://www.phy.olemiss.edu/~luca/Topics/qm/collapse.html>)
- [25] Greenstein, George; Zajonc, Arthur (2006). *The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics, Second edition* (<http://books.google.com/books?id=5t0tm0FB1CsC&pg=PA215>). Jones and Bartlett Publishers, Inc. p. 215. ISBN 0-7637-2470-X.

- ..
- [26] Farside.ph.utexas.edu (<http://farside.ph.utexas.edu/teaching/qmech/lectures/node28.html>)
 - [27] Mathews, Piravonu Mathews; Venkatesan, K. (1976). *A Textbook of Quantum Mechanics* (http://books.google.com/books?id=_qzs1DD3TcsC&pg=PA36). Tata McGraw-Hill. p. 36. ISBN 0-07-096510-2. .
 - [28] Physics.ukzn.ac.za (<http://physics.ukzn.ac.za/~petruccione/Phys120/Wave Functions and the Schrödinger Equation.pdf>)
 - [29] Reddit.com (http://www.reddit.com/r/philosophy/comments/8p2qv/determinism_and_naive_realism/)
 - [30] <http://demonstrations.wolfram.com/TimeEvolutionOfAWavepacketInASquareWell/>
 - [31] Greenstein, George; Zajonc, Arthur (2006). *The Quantum Challenge: Modern Research on the Foundations of Quantum Mechanics, Second edition* (<http://books.google.com/books?id=5t0tm0FB1CsC&pg=PA215>). Jones and Bartlett Publishers, Inc. p. 215. ISBN 0-7637-2470-X.
 - ..
 - [32] P.A.M. Dirac, *The Principles of Quantum Mechanics*, Clarendon Press, Oxford, 1930.
 - [33] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932 (English translation: *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955).
 - [34] "The Nobel Prize in Physics 1979" (http://nobelprize.org/nobel_prizes/physics/laureates/1979/index.html). Nobel Foundation. . Retrieved 2010-02-16.
 - [35] Complex Elliptic Pendulum (<http://arxiv.org/abs/1001.0131>), Carl M. Bender, Daniel W. Hook, Karta Kooner
 - [36] Derivation of particle in a box, chemistry.tidalswan.com (http://chemistry.tidalswan.com/index.php?title=Quantum_Mechanics)
 - [37] Life on the lattice: The most accurate theory we have. (<http://latticeqcd.blogspot.com/2005/06/most-accurate-theory-we-have.html>)
 - [38] Parker, B. (1993). *Overcoming some of the problems*. pp. 259–279.
 - [39] "There is as yet no logically consistent and complete relativistic quantum field theory.", p. 4. — V. B. Berestetskii, E. M. Lifshitz, L P Pitaevskii (1971). J. B. Sykes, J. S. Bell (translators). *Relativistic Quantum Theory 4, part I. Course of Theoretical Physics (Landau and Lifshitz)* ISBN 0 08 016025 5
 - [40] Books.google.com (http://books.google.com/books?id=vdXU6SD4_UYC)
 - [41] en.wikibooks.org (http://en.wikibooks.org/wiki/Computational_chemistry/Applications_of_molecular_quantum_mechanics)
 - [42] Discovermagazine.com (http://discovermagazine.com/2009/feb/13-is-quantum-mechanics-controlling-your-thoughts/article_view?b_start:int=1&-C)
 - [43] Plato.stanford.edu (<http://plato.stanford.edu/entries/qm-action-distance/>)
 - [44] Plato.stanford.edu (<http://plato.stanford.edu/entries/qm-everett/>)
 - [45] www-physics.lbl.gov (<http://www-physics.lbl.gov/~stapp/PTRS.pdf>)
 - [46] <http://www.physik.fu-berlin.de/~kleinert/b5>
 - [47] http://www.lightandmatter.com/html_books/6mr/ch01/ch01.html
 - [48] http://www-history.mcs.st-andrews.ac.uk/history/HistTopics/The_Quantum_age_begins.html
 - [49] http://www.quantiki.org/wiki/index.php/Introduction_to_Quantum_Theory
 - [50] <http://bethe.cornell.edu/>
 - [51] <http://www.nonlocal.com/hbar/>
 - [52] <http://www.freebookcentre.net/Physics/Quantum-Mechanics-Books.html>
 - [53] <http://arxiv.org/abs/quant-ph/0605180>
 - [54] <http://ocw.mit.edu/OcwWeb/Chemistry/index.htm>
 - [55] <http://ocw.mit.edu/OcwWeb/Chemistry/5-61Fall-2004/CourseHome/index.htm>
 - [56] <http://ocw.mit.edu/OcwWeb/Chemistry/5-73Fall-2005/CourseHome/index.htm>
 - [57] <http://ocw.mit.edu/OcwWeb/Chemistry/5-74Spring-2005/CourseHome/index.htm>
 - [58] <http://ocw.mit.edu/OcwWeb/Physics/index.htm>
 - [59] <http://ocw.mit.edu/OcwWeb/Physics/8-04Spring-2006/CourseHome/index.htm>
 - [60] <http://ocw.mit.edu/OcwWeb/Physics/8-05Fall-2004/CourseHome/index.htm>
 - [61] <http://ocw.mit.edu/OcwWeb/Physics/8-06Spring-2005/CourseHome/index.htm>
 - [62] <http://www.youtube.com/stanford#g/c/84C10A9CB1D13841>
 - [63] http://continuingstudies.stanford.edu/courses/course.php?cid=20072_PHY%2025
 - [64] <http://www.physics.csbsju.edu/QM/>
 - [65] <http://www.imperial.ac.uk/quantuminformation/qi/tutorials>
 - [66] <http://www.sparknotes.com/testprep/books/sat2/physics/chapter19section3.rhtml>
 - [67] <http://www.quantum-physics.polytechnique.fr>
 - [68] <http://www.quantumlab.de>
 - [69] <http://www.motionmountain.net/download.html>
 - [70] <http://www.nanohub.org/topics/AQME>
 - [71] <http://www.lsr.ph.ic.ac.uk/~plenio/lecture.pdf>
 - [72] <http://farside.ph.utexas.edu/teaching/qm/389.pdf>
 - [73] <http://www.hedweb.com/manworld.htm>
 - [74] <http://www.mtnmath.com/faq/meas-qm.html>
 - [75] <http://www.newscientist.com/channel/fundamentals/quantum-world>

[76] http://www.sciencedaily.com/news/matter_energy/quantum_physics/

[77] <http://www.nytimes.com/2005/12/27/science/27eins.html?ex=1293339600&en=caf5d835203c3500&ei=5090>

[78] <http://plato.stanford.edu/entries/qm>

Quantum field theory

Quantum field theory (QFT)^[1] provides a theoretical framework for constructing quantum mechanical models of systems classically described by fields or (especially in a condensed matter context) many-body systems. It is widely used in particle physics and condensed matter physics. Most theories in modern particle physics, including the Standard Model of elementary particles and their interactions, are formulated as relativistic quantum field theories. Quantum field theories are used in many circumstances, especially those where the number of particles fluctuates—for example, in the BCS theory of superconductivity.

In perturbative quantum field theory, the forces between particles are mediated by other particles. The electromagnetic force between two electrons is caused by an exchange of photons. Intermediate vector bosons mediate the weak force and gluons mediate the strong force. There is currently no complete quantum theory of the remaining fundamental force, gravity, but many of the proposed theories postulate the existence of a graviton particle which mediates it. These force-carrying particles are virtual particles and, by definition, cannot be detected while carrying the force, because such detection will imply that the force is not being carried. In addition, the notion of "force mediating particle" comes from perturbation theory, and thus does not make sense in a context of bound states.

In QFT photons are not thought of as 'little billiard balls', they are considered to be field quanta - necessarily chunked ripples in a field that 'look like' particles. Fermions, like the electron, can also be described as ripples in a field, where each kind of fermion has its own field. In summary, the classical visualisation of "everything is particles and fields", in quantum field theory, resolves into "everything is particles", which then resolves into "everything is fields". In the end, particles are regarded as excited states of a field (field quanta).

History

Quantum field theory originated in the 1920s from the problem of creating a quantum mechanical theory of the electromagnetic field. In 1926, Max Born, Pascual Jordan, and Werner Heisenberg constructed such a theory by expressing the field's internal degrees of freedom as an infinite set of harmonic oscillators and by employing the usual procedure for quantizing those oscillators (canonical quantization). Max Planck, a physicist at the University of Kiel, observed the behavior at the atomic level of radiation and heat on matter. He observed that the energy absorbed or emitted was contained in small, discrete (i.e. individual) energy packets called quanta. This theory assumed that no electric charges or currents were present, and today would be called a free field theory. The first reasonably complete theory of quantum electrodynamics, which included both the electromagnetic field and electrically charged matter (specifically, electrons) as quantum mechanical objects, was created by Paul Dirac in 1927.^[2] This quantum field theory could be used to model important processes such as the emission of a photon by an electron dropping into a quantum state of lower energy, a process in which the *number of particles changes* — one atom in the initial state becomes an atom plus a photon in the final state. It is now understood that the ability to describe such processes is one of the most important features of quantum field theory.

It was evident from the beginning that a proper quantum treatment of the electromagnetic field had to somehow incorporate Einstein's relativity theory, which had after all grown out of the study of classical electromagnetism. This need to *put together relativity and quantum mechanics* was the second major motivation in the development of quantum field theory. Pascual Jordan and Wolfgang Pauli showed in 1928 that quantum fields could be made to behave in the way predicted by special relativity during coordinate transformations (specifically, they showed that the field commutators were Lorentz invariant), and in 1933 Niels Bohr and Leon Rosenfeld showed that this result

could be interpreted as a limitation on the ability to measure fields at space-like separations, exactly as required by relativity. A further boost for quantum field theory came with the discovery of the Dirac equation, a single-particle equation obeying both relativity and quantum mechanics, when it was shown that several of its undesirable properties (such as negative-energy states) could be eliminated by reformulating the Dirac equation as a quantum field theory. This work was performed by Wendell Furry, Robert Oppenheimer, Vladimir Fock, and others.

The third thread in the development of quantum field theory was the need to *handle the statistics of many-particle systems* consistently and with ease. In 1927, Jordan tried to extend the canonical quantization of fields to the many-body wavefunctions of identical particles, a procedure that is sometimes called second quantization. In 1928, Jordan and Eugene Wigner found that the quantum field describing electrons, or other fermions, had to be expanded using anti-commuting creation and annihilation operators due to the Pauli exclusion principle. This thread of development was incorporated into many-body theory, and strongly influenced condensed matter physics and nuclear physics.

Despite its early successes, quantum field theory was plagued by several serious theoretical difficulties. Many seemingly-innocuous physical quantities, such as the energy shift of electron states due to the presence of the electromagnetic field, gave infinity — a nonsensical result — when computed using quantum field theory. This "divergence problem" was solved during the 1940s by Bethe, Tomonaga, Schwinger, Feynman, and Dyson, through the procedure known as renormalization. This phase of development culminated with the construction of the modern theory of quantum electrodynamics (QED). Beginning in the 1950s with the work of Yang and Mills, QED was generalized to a class of quantum field theories known as gauge theories. The 1960s and 1970s saw the formulation of a gauge theory now known as the Standard Model of particle physics, which describes all known elementary particles and the interactions between them. The weak interaction part of the standard model was formulated by Sheldon Glashow, with the Higgs mechanism added by Steven Weinberg and Abdus Salam. The theory was shown to be renormalizable and hence consistent by Gerardus 't Hooft and Martinus Veltman.

Also during the 1970s, parallel developments in the study of phase transitions in condensed matter physics led Leo Kadanoff, Michael Fisher and Kenneth Wilson (extending work of Ernst Stueckelberg, Andre Peterman, Murray Gell-Mann and Francis Low) to a set of ideas and methods known as the renormalization group. By providing a better physical understanding of the renormalization procedure invented in the 1940s, the renormalization group sparked what has been called the "grand synthesis" of theoretical physics, uniting the quantum field theoretical techniques used in particle physics and condensed matter physics into a single theoretical framework.

The study of quantum field theory is alive and flourishing, as are applications of this method to many physical problems. It remains one of the most vital areas of theoretical physics today, providing a common language to many branches of physics.

Principles of quantum field theory

Classical fields and quantum fields

Quantum mechanics, in its most general formulation, is a theory of abstract operators (observables) acting on an abstract state space (Hilbert space), where the observables represent physically-observable quantities and the state space represents the possible states of the system under study. Furthermore, each observable corresponds, in a technical sense, to the classical idea of a degree of freedom. For instance, the fundamental observables associated with the motion of a single quantum mechanical particle are the position and momentum operators $\hat{x}$ and $\hat{p}$.

Ordinary quantum mechanics deals with systems such as this, which possess a small set of degrees of freedom.

(It is important to note, at this point, that this article does not use the word "particle" in the context of wave–particle duality. In quantum field theory, "particle" is a generic term for any discrete quantum mechanical entity, such as an electron or photon, which can behave like classical particles or classical waves under different experimental conditions.)

A **quantum field** is a quantum mechanical system containing a large, and possibly infinite, number of degrees of freedom. This is not as exotic a situation as one might think (e.g., in an infinite-dimensional vector space, a vector is just a function as ordinary as $f(x) = x^2$; which is, of course, familiar territory). A classical field contains a set of degrees of freedom at each point of space; for instance, the classical electromagnetic field defines two vectors — the electric field and the magnetic field — that can in principle take on distinct values for each position r . When the field *as a whole* is considered as a quantum mechanical system, its observables form an infinite (in fact uncountable) set, because r is continuous.

Furthermore, the degrees of freedom in a quantum field are arranged in "repeated" sets. For example, the degrees of freedom in an electromagnetic field can be grouped according to the position r , with exactly two vectors for each r . Note that r is an ordinary number that "indexes" the observables; it is not to be confused with the position operator $\hat{x}$ encountered in ordinary quantum mechanics, which is an observable. (Thus, ordinary quantum mechanics is sometimes referred to as "zero-dimensional quantum field theory", because it contains only a single set of observables.)

It is also important to note that there is nothing special about r because, as it turns out, there is generally more than one way of indexing the degrees of freedom in the field.

In the following sections, we will show how these ideas can be used to construct a quantum mechanical theory with the desired properties. We will begin by discussing single-particle quantum mechanics and the associated theory of many-particle quantum mechanics. Then, by finding a way to index the degrees of freedom in the many-particle problem, we will construct a quantum field and study its implications.

Single-particle and many-particle quantum mechanics

In ordinary quantum mechanics, the time-dependent one-dimensional Schrödinger equation describing the time evolution of the quantum state of a single non-relativistic particle is

$$\left[\frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V(\mathbf{x}) \right] |\psi(t)\rangle = i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle,$$

where m is the particle's mass, V is the applied potential, and $|\psi\rangle$ denotes the quantum state (we are using bra-ket notation).

We wish to consider how this problem generalizes to N particles. There are two motivations for studying the many-particle problem. The first is a straightforward need in condensed matter physics, where typically the number of particles is on the order of Avogadro's number (6.0221415×10^{23}). The second motivation for the many-particle problem arises from particle physics and the desire to incorporate the effects of special relativity. If one attempts to include the relativistic rest energy into the above equation (in quantum mechanics where position is an observable), the result is either the Klein-Gordon equation or the Dirac equation. However, these equations have many unsatisfactory qualities; for instance, they possess energy eigenvalues which extend to $-\infty$, so that there seems to be no easy definition of a ground state. It turns out that such inconsistencies arise from relativistic wavefunctions having a probabilistic interpretation in position space, as probability conservation is not a relativistically covariant concept. In quantum field theory, unlike in quantum mechanics, position is not an observable, and thus, one does not need the concept of a position-space probability density. For quantum fields whose interaction can be treated perturbatively, this is equivalent to neglecting the possibility of dynamically creating or destroying particles, which is a crucial aspect of relativistic quantum theory. Einstein's famous mass-energy relation allows for the possibility that sufficiently massive particles can decay into several lighter particles, and sufficiently energetic particles can combine to form massive particles. For example, an electron and a positron can annihilate each other to create photons. This suggests that a consistent relativistic quantum theory should be able to describe many-particle dynamics.

Furthermore, we will assume that the N particles are indistinguishable. As described in the article on identical particles, this implies that the state of the entire system must be either symmetric (bosons) or antisymmetric

(fermions) when the coordinates of its constituent particles are exchanged. These multi-particle states are rather complicated to write. For example, the general quantum state of a system of N bosons is written as

$$|\phi_1 \cdots \phi_N\rangle = \sqrt{\frac{\prod_j N_j!}{N!}} \sum_{p \in S_N} |\phi_{p(1)}\rangle \cdots |\phi_{p(N)}\rangle,$$

where $|\phi_i\rangle$ are the single-particle states, N_j is the number of particles occupying state j , and the sum is taken over all possible permutations p acting on N elements. In general, this is a sum of $N!(N \text{ factorial})$ distinct terms, which quickly becomes unmanageable as N increases. The way to simplify this problem is to turn it into a quantum field theory.

Second quantization

In this section, we will describe a method for constructing a quantum field theory called **second quantization**. This basically involves choosing a way to index the quantum mechanical degrees of freedom in the space of multiple identical-particle states. It is based on the Hamiltonian formulation of quantum mechanics; several other approaches exist, such as the Feynman path integral,^[3] which uses a Lagrangian formulation. For an overview, see the article on quantization.

Second quantization of bosons

For simplicity, we will first discuss second quantization for bosons, which form perfectly symmetric quantum states. Let us denote the mutually orthogonal single-particle states by $|\phi_1\rangle, |\phi_2\rangle, |\phi_3\rangle$, and so on. For example, the 3-particle state with one particle in state $|\phi_1\rangle$ and two in state $|\phi_2\rangle$ is

$$\frac{1}{\sqrt{3}} [|\phi_1\rangle|\phi_2\rangle|\phi_2\rangle + |\phi_2\rangle|\phi_1\rangle|\phi_2\rangle + |\phi_2\rangle|\phi_2\rangle|\phi_1\rangle].$$

The first step in second quantization is to express such quantum states in terms of **occupation numbers**, by listing the number of particles occupying each of the single-particle states $|\phi_1\rangle, |\phi_2\rangle$, etc. This is simply another way of labelling the states. For instance, the above 3-particle state is denoted as

$$|1, 2, 0, 0, \dots\rangle.$$

The next step is to expand the N -particle state space to include the state spaces for all possible values of N . This extended state space, known as a Fock space, is composed of the state space of a system with no particles (the so-called vacuum state), plus the state space of a 1-particle system, plus the state space of a 2-particle system, and so forth. It is easy to see that there is a one-to-one correspondence between the occupation number representation and valid boson states in the Fock space.

At this point, the quantum mechanical system has become a quantum field in the sense we described above. The field's elementary degrees of freedom are the occupation numbers, and each occupation number is indexed by a number $j \cdots$, indicating which of the single-particle states $|\phi_1\rangle, |\phi_2\rangle, \dots |\phi_j\rangle \cdots$ it refers to.

The properties of this quantum field can be explored by defining creation and annihilation operators, which add and subtract particles. They are analogous to "ladder operators" in the quantum harmonic oscillator problem, which added and subtracted energy quanta. However, these operators literally create and annihilate particles of a given quantum state. The bosonic annihilation operator a_2 and creation operator $a_2^\dagger$ have the following effects:

$$\begin{aligned} a_2 |N_1, N_2, N_3, \dots\rangle &= \sqrt{N_2} |N_1, (N_2 - 1), N_3, \dots\rangle, \\ a_2^\dagger |N_1, N_2, N_3, \dots\rangle &= \sqrt{N_2 + 1} |N_1, (N_2 + 1), N_3, \dots\rangle. \end{aligned}$$

It can be shown that these are operators in the usual quantum mechanical sense, i.e. linear operators acting on the Fock space. Furthermore, they are indeed Hermitian conjugates, which justifies the way we have written them. They can be shown to obey the commutation relation

$$[a_i, a_j] = 0 \quad , \quad [a_i^\dagger, a_j^\dagger] = 0 \quad , \quad [a_i, a_j^\dagger] = \delta_{ij},$$

where δ stands for the Kronecker delta. These are precisely the relations obeyed by the ladder operators for an infinite set of independent quantum harmonic oscillators, one for each single-particle state. Adding or removing bosons from each state is therefore analogous to exciting or de-exciting a quantum of energy in a harmonic oscillator.

The Hamiltonian of the quantum field (which, through the Schrödinger equation, determines its dynamics) can be written in terms of creation and annihilation operators. For instance, the Hamiltonian of a field of free (non-interacting) bosons is

$$H = \sum_k E_k a_k^\dagger a_k,$$

where E_k is the energy of the k -th single-particle energy eigenstate. Note that

$$a_k^\dagger a_k |\dots, N_k, \dots\rangle = N_k |\dots, N_k, \dots\rangle.$$

Second quantization of fermions

It turns out that a different definition of creation and annihilation must be used for describing fermions. According to the Pauli exclusion principle, fermions cannot share quantum states, so their occupation numbers N_i can only take on the value 0 or 1. The fermionic annihilation operators c and creation operators $c^\dagger$ are defined by their actions on a Fock state thus

$$\begin{aligned} c_j |N_1, N_2, \dots, N_j = 0, \dots\rangle &= 0 \\ c_j |N_1, N_2, \dots, N_j = 1, \dots\rangle &= (-1)^{(N_1 + \dots + N_{j-1})} |N_1, N_2, \dots, N_j = 0, \dots\rangle \\ c_j^\dagger |N_1, N_2, \dots, N_j = 0, \dots\rangle &= (-1)^{(N_1 + \dots + N_{j-1})} |N_1, N_2, \dots, N_j = 1, \dots\rangle \\ c_j^\dagger |N_1, N_2, \dots, N_j = 1, \dots\rangle &= 0. \end{aligned}$$

These obey an anticommutation relation:

$$\{c_i, c_j\} = 0 \quad , \quad \{c_i^\dagger, c_j^\dagger\} = 0 \quad , \quad \{c_i, c_j^\dagger\} = \delta_{ij}.$$

One may notice from this that applying a fermionic creation operator twice gives zero, so it is impossible for the particles to share single-particle states, in accordance with the exclusion principle.

Field operators

We have previously mentioned that there can be more than one way of indexing the degrees of freedom in a quantum field. Second quantization indexes the field by enumerating the single-particle quantum states. However, as we have discussed, it is more natural to think about a "field", such as the electromagnetic field, as a set of degrees of freedom indexed by position.

To this end, we can define *field operators* that create or destroy a particle at a particular point in space. In particle physics, these operators turn out to be more convenient to work with, because they make it easier to formulate theories that satisfy the demands of relativity.

Single-particle states are usually enumerated in terms of their momenta (as in the particle in a box problem.) We can construct field operators by applying the Fourier transform to the creation and annihilation operators for these states. For example, the bosonic field annihilation operator $\phi(\mathbf{r})$ is

$$\phi(\mathbf{r}) \stackrel{\text{def}}{=} \sum_j e^{i\mathbf{k}_j \cdot \mathbf{r}} a_j.$$

The bosonic field operators obey the commutation relation

$$[\phi(\mathbf{r}), \phi(\mathbf{r}')] = 0 \quad , \quad [\phi^\dagger(\mathbf{r}), \phi^\dagger(\mathbf{r}')] = 0 \quad , \quad [\phi(\mathbf{r}), \phi^\dagger(\mathbf{r}')] = \delta^3(\mathbf{r} - \mathbf{r}')$$

where $\delta(x)$ stands for the Dirac delta function. As before, the fermionic relations are the same, with the commutators replaced by anticommutators.

It should be emphasized that the field operator is *not* the same thing as a single-particle wavefunction. The former is an operator acting on the Fock space, and the latter is just a scalar field. However, they are closely related, and are indeed commonly denoted with the same symbol. If we have a Hamiltonian with a space representation, say

$$H = -\frac{\hbar^2}{2m} \sum_i \nabla_i^2 + \sum_{i < j} U(|\mathbf{r}_i - \mathbf{r}_j|)$$

where the indices i and j run over all particles, then the field theory Hamiltonian is

$$H = -\frac{\hbar^2}{2m} \int d^3r \phi^\dagger(\mathbf{r}) \nabla^2 \phi(\mathbf{r}) + \int d^3r \int d^3r' \phi^\dagger(\mathbf{r}) \phi^\dagger(\mathbf{r}') U(|\mathbf{r} - \mathbf{r}'|) \phi(\mathbf{r}') \phi(\mathbf{r}).$$

This looks remarkably like an expression for the expectation value of the energy, with ϕ playing the role of the wavefunction. This relationship between the field operators and wavefunctions makes it very easy to formulate field theories starting from space-projected Hamiltonians.

Implications of quantum field theory

Unification of fields and particles

The "second quantization" procedure that we have outlined in the previous section takes a set of single-particle quantum states as a starting point. Sometimes, it is impossible to define such single-particle states, and one must proceed directly to quantum field theory. For example, a quantum theory of the electromagnetic field *must* be a quantum field theory, because it is impossible (for various reasons) to define a wavefunction for a single photon. In such situations, the quantum field theory can be constructed by examining the mechanical properties of the classical field and guessing the corresponding quantum theory. For free (non-interacting) quantum fields, the quantum field theories obtained in this way have the same properties as those obtained using second quantization, such as well-defined creation and annihilation operators obeying commutation or anticommutation relations.

Quantum field theory thus provides a unified framework for describing "field-like" objects (such as the electromagnetic field, whose excitations are photons) and "particle-like" objects (such as electrons, which are treated as excitations of an underlying electron field), so long as one can treat interactions as "perturbations" of free fields. There are still unsolved problems relating to the more general case of interacting fields which may or may not be adequately described by perturbation theory. For more on this topic, see Haag's theorem.

Physical meaning of particle indistinguishability

The second quantization procedure relies crucially on the particles being identical. We would not have been able to construct a quantum field theory from a distinguishable many-particle system, because there would have been no way of separating and indexing the degrees of freedom.

Many physicists prefer to take the converse interpretation, which is that *quantum field theory explains what identical particles are*. In ordinary quantum mechanics, there is not much theoretical motivation for using symmetric (bosonic) or antisymmetric (fermionic) states, and the need for such states is simply regarded as an empirical fact. From the point of view of quantum field theory, particles are identical if and only if they are excitations of the same underlying quantum field. Thus, the question "why are all electrons identical?" arises from mistakenly regarding individual electrons as fundamental objects, when in fact it is only the electron field that is fundamental.

Particle conservation and non-conservation

During second quantization, we started with a Hamiltonian and state space describing a fixed number of particles (N), and ended with a Hamiltonian and state space for an arbitrary number of particles. Of course, in many common situations N is an important and perfectly well-defined quantity, e.g. if we are describing a gas of atoms sealed in a box. From the point of view of quantum field theory, such situations are described by quantum states that are eigenstates of the number operator $\hat{N}$, which measures the total number of particles present. As with any quantum mechanical observable, $\hat{N}$ is conserved if it commutes with the Hamiltonian. In that case, the quantum state is trapped in the N -particle subspace of the total Fock space, and the situation could equally well be described by ordinary N -particle quantum mechanics.

For example, we can see that the free-boson Hamiltonian described above conserves particle number. Whenever the Hamiltonian operates on a state, each particle destroyed by an annihilation operator a_k is immediately put back by the creation operator $a_k^\dagger$.

On the other hand, it is possible, and indeed common, to encounter quantum states that are *not* eigenstates of $\hat{N}$, which do not have well-defined particle numbers. Such states are difficult or impossible to handle using ordinary quantum mechanics, but they can be easily described in quantum field theory as quantum superpositions of states having different values of N . For example, suppose we have a bosonic field whose particles can be created or destroyed by interactions with a fermionic field. The Hamiltonian of the combined system would be given by the Hamiltonians of the free boson and free fermion fields, plus a "potential energy" term such as

$$H_I = \sum_{k,q} V_q (a_q + a_{-q}^\dagger) c_{k+q}^\dagger c_k,$$

where $a_k^\dagger$ and a_k denotes the bosonic creation and annihilation operators, $c_k^\dagger$ and c_k denotes the fermionic creation and annihilation operators, and V_q is a parameter that describes the strength of the interaction. This "interaction term" describes processes in which a fermion in state k either absorbs or emits a boson, thereby being kicked into a different eigenstate $k + q$. (In fact, this type of Hamiltonian is used to describe interaction between conduction electrons and phonons in metals. The interaction between electrons and photons is treated in a similar way, but is a little more complicated because the role of spin must be taken into account.) One thing to notice here is that even if we start out with a fixed number of bosons, we will typically end up with a superposition of states with different numbers of bosons at later times. The number of fermions, however, is conserved in this case. In condensed matter physics, states with ill-defined particle numbers are particularly important for describing the various superfluids. Many of the defining characteristics of a superfluid arise from the notion that its quantum state is a superposition of states with different particle numbers. In addition, the concept of a coherent state (used to model the laser and the BCS ground state) refers to a state with an ill-defined particle number but a well-defined phase.

Axiomatic approaches

The preceding description of quantum field theory follows the spirit in which most physicists approach the subject. However, it is not mathematically rigorous. Over the past several decades, there have been many attempts to put quantum field theory on a firm mathematical footing by formulating a set of axioms for it. These attempts fall into two broad classes.

The first class of axioms, first proposed during the 1950s, include the Wightman, Osterwalder-Schrader, and Haag-Kastler systems. They attempted to formalize the physicists' notion of an "operator-valued field" within the context of functional analysis, and enjoyed limited success. It was possible to prove that any quantum field theory satisfying these axioms satisfied certain general theorems, such as the spin-statistics theorem and the CPT theorem. Unfortunately, it proved extraordinarily difficult to show that any realistic field theory, including the Standard Model, satisfied these axioms. Most of the theories that could be treated with these analytic axioms were physically trivial, being restricted to low-dimensions and lacking interesting dynamics. The construction of theories satisfying one of these sets of axioms falls in the field of constructive quantum field theory. Important work was done in this

area in the 1970s by Segal, Glimm, Jaffe and others.

During the 1980s, a second set of axioms based on geometric ideas was proposed. This line of investigation, which restricts its attention to a particular class of quantum field theories known as topological quantum field theories, is associated most closely with Michael Atiyah and Graeme Segal, and was notably expanded upon by Edward Witten, Richard Borcherds, and Maxim Kontsevich. However, most physically-relevant quantum field theories, such as the Standard Model, are not topological quantum field theories; the quantum field theory of the fractional quantum Hall effect is a notable exception. The main impact of axiomatic topological quantum field theory has been on mathematics, with important applications in representation theory, algebraic topology, and differential geometry.

Finding the proper axioms for quantum field theory is still an open and difficult problem in mathematics. One of the Millennium Prize Problems—proving the existence of a mass gap in Yang-Mills theory—is linked to this issue.

Phenomena associated with quantum field theory

In the previous part of the article, we described the most general properties of quantum field theories. Some of the quantum field theories studied in various fields of theoretical physics possess additional special properties, such as renormalizability, gauge symmetry, and supersymmetry. These are described in the following sections.

Renormalization

Early in the history of quantum field theory, it was found that many seemingly innocuous calculations, such as the perturbative shift in the energy of an electron due to the presence of the electromagnetic field, give infinite results. The reason is that the perturbation theory for the shift in an energy involves a sum over all other energy levels, and there are infinitely many levels at short distances which each give a finite contribution.

Many of these problems are related to failures in classical electrodynamics that were identified but unsolved in the 19th century, and they basically stem from the fact that many of the supposedly "intrinsic" properties of an electron are tied to the electromagnetic field which it carries around with it. The energy carried by a single electron—its self energy—is not simply the bare value, but also includes the energy contained in its electromagnetic field, its attendant cloud of photons. The energy in a field of a spherical source diverges in both classical and quantum mechanics, but as discovered by Weisskopf, in quantum mechanics the divergence is much milder, going only as the logarithm of the radius of the sphere.

The solution to the problem, presciently suggested by Stueckelberg, independently by Bethe after the crucial experiment by Lamb, implemented at one loop by Schwinger, and systematically extended to all loops by Feynman and Dyson, with converging work by Tomonaga in isolated postwar Japan, is called renormalization. The technique of renormalization recognizes that the problem is essentially purely mathematical, that extremely short distances are at fault. In order to define a theory on a continuum, first place a cutoff on the fields, by postulating that quanta cannot have energies above some extremely high value. This has the effect of replacing continuous space by a structure where very short wavelengths do not exist, as on a lattice. Lattices break rotational symmetry, and one of the crucial contributions made by Feynman, Pauli and Villars, and modernized by 't Hooft and Veltman, is a symmetry preserving cutoff for perturbation theory. There is no known symmetrical cutoff outside of perturbation theory, so for rigorous or numerical work people often use an actual lattice.

On a lattice, every quantity is finite but depends on the spacing. When taking the limit of zero spacing, we make sure that the physically-observable quantities like the observed electron mass stay fixed, which means that the constants in the Lagrangian defining the theory depend on the spacing. Hopefully, by allowing the constants to vary with the lattice spacing, all the results at long distances become insensitive to the lattice, defining a continuum limit.

The renormalization procedure only works for a certain class of quantum field theories, called **renormalizable quantum field theories**. A theory is **perturbatively renormalizable** when the constants in the Lagrangian only diverge at worst as logarithms of the lattice spacing for very short spacings. The continuum limit is then well defined

in perturbation theory, and even if it is not fully well defined non-perturbatively, the problems only show up at distance scales which are exponentially small in the inverse coupling for weak couplings. The Standard Model of particle physics is perturbatively renormalizable, and so are its component theories (quantum electrodynamics/electroweak theory and quantum chromodynamics). Of the three components, quantum electrodynamics is believed to not have a continuum limit, while the asymptotically free $SU(2)$ and $SU(3)$ weak hypercharge and strong color interactions are nonperturbatively well defined.

The renormalization group describes how renormalizable theories emerge as the long distance low-energy effective field theory for any given high-energy theory. Because of this, renormalizable theories are insensitive to the precise nature of the underlying high-energy short-distance phenomena. This is a blessing because it allows physicists to formulate low energy theories without knowing the details of high energy phenomenon. It is also a curse, because once a renormalizable theory like the standard model is found to work, it gives very few clues to higher energy processes. The only way high energy processes can be seen in the standard model is when they allow otherwise forbidden events, or if they predict quantitative relations between the coupling constants.

Gauge freedom

A gauge theory is a theory that admits a symmetry with a local parameter. For example, in every quantum theory the global phase of the wave function is arbitrary and does not represent something physical. Consequently, the theory is invariant under a global change of phases (adding a constant to the phase of all wave functions, everywhere); this is a global symmetry. In quantum electrodynamics, the theory is also invariant under a *local* change of phase, that is - one may shift the phase of all wave functions so that the shift may be different at every point in space-time. This is a *local* symmetry. However, in order for a well-defined derivative operator to exist, one must introduce a new field, the gauge field, which also transforms in order for the local change of variables (the phase in our example) not to affect the derivative. In quantum electrodynamics this gauge field is the electromagnetic field. The change of local gauge of variables is termed gauge transformation.

In quantum field theory the excitations of fields represent particles. The particle associated with excitations of the gauge field is the gauge boson, which is the photon in the case of quantum electrodynamics.

The degrees of freedom in quantum field theory are local fluctuations of the fields. The existence of a gauge symmetry reduces the number of degrees of freedom, simply because some fluctuations of the fields can be transformed to zero by gauge transformations, so they are equivalent to having no fluctuations at all, and they therefore have no physical meaning. Such fluctuations are usually called "non-physical degrees of freedom" or *gauge artifacts*; usually some of them have a negative norm, making them inadequate for a consistent theory. Therefore, if a classical field theory has a gauge symmetry, then its quantized version (i.e. the corresponding quantum field theory) will have this symmetry as well. In other words, a gauge symmetry cannot have a quantum anomaly. If a gauge symmetry is anomalous (i.e. not kept in the quantum theory) then the theory is non-consistent: for example, in quantum electrodynamics, had there been a gauge anomaly, this would require the appearance of photons with longitudinal polarization and polarization in the time direction, the latter having a negative norm, rendering the theory inconsistent; another possibility would be for these photons to appear only in intermediate processes but not in the final products of any interaction, making the theory non unitary and again inconsistent (see optical theorem).

In general, the gauge transformations of a theory consist of several different transformations, which may not be commutative. These transformations are together described by a mathematical object known as a gauge group. Infinitesimal gauge transformations are the gauge group generators. Therefore the number of gauge bosons is the group dimension (i.e. number of generators forming a basis).

All the fundamental interactions in nature are described by gauge theories. These are:

- Quantum electrodynamics, whose gauge transformation is a local change of phase, so that the gauge group is $U(1)$. The gauge boson is the photon.
- Quantum chromodynamics, whose gauge group is $SU(3)$. The gauge bosons are eight gluons.

- The electroweak Theory, whose gauge group is $U(1) \otimes SU(2)$ (a direct product of $U(1)$ and $SU(2)$).
- Gravity, whose classical theory is general relativity, admits the equivalence principle which is a form of gauge symmetry, however it is explicitly non-renormalizable.

Supersymmetry

Supersymmetry assumes that every fundamental fermion has a superpartner that is a boson and vice versa. It was introduced in order to solve the so-called Hierarchy Problem, that is, to explain why particles not protected by any symmetry (like the Higgs boson) do not receive radiative corrections to its mass driving it to the larger scales (GUT, Planck...). It was soon realized that supersymmetry has other interesting properties: its gauged version is an extension of general relativity (Supergravity), and it is a key ingredient for the consistency of string theory.

The way supersymmetry protects the hierarchies is the following: since for every particle there is a superpartner with the same mass, any loop in a radiative correction is cancelled by the loop corresponding to its superpartner, rendering the theory UV finite.

Since no superpartners have yet been observed, if supersymmetry exists it must be broken (through a so-called soft term, which breaks supersymmetry without ruining its helpful features). The simplest models of this breaking require that the energy of the superpartners not be too high; in these cases, supersymmetry is expected to be observed by experiments at the Large Hadron Collider.

See also

- List of quantum field theories
- Feynman path integral
- Quantum chromodynamics
- Quantum electrodynamics
- Quantum flavordynamics
- Quantum geometrodynamics
- Quantum hydrodynamics
- Quantum magnetodynamics
- Quantum triviality
- Schwinger-Dyson equation
- Relation between Schrödinger's equation and the path integral formulation of quantum mechanics
- Basic concepts of quantum mechanics
- Relationship between string theory and quantum field theory
- Abraham-Lorentz force
- Form factor
- Photon polarization
- Theoretical and experimental justification for the Schrödinger equation
- Invariance mechanics
- Green–Kubo relations
- Green's function (many-body theory)
- Common integrals in quantum field theory

Further reading

General readers:

- Feynman, R.P. (2001) [1964]. *The Character of Physical Law*. MIT Press. ISBN 0262560038.
- Feynman, R.P. (2006) [1985]. *QED: The Strange Theory of Light and Matter*. Princeton University Press. ISBN 0691125759.
- Gribbin, J. (1998). *Q is for Quantum: Particle Physics from A to Z*. Weidenfeld & Nicolson. ISBN 0297817523.
- Schumm, Bruce A. (2004) *Deep Down Things*. Johns Hopkins Univ. Press. Chpt. 4.

Introductory texts:

- Bogoliubov, N.; Shirkov, D. (1982). *Quantum Fields*. Benjamin-Cummings. ISBN 0805309837.
- Frampton, P.H. (2000). *Gauge Field Theories. Frontiers in Physics (2nd ed.)*. Wiley.
- Greiner, W; Müller, B. (2000). *Gauge Theory of Weak Interactions*. Springer. ISBN 3-540-67672-4.

- Itzykson, C.; Zuber, J.-B. (1980). *Quantum Field Theory*. McGraw-Hill. ISBN 0-07-032071-3.
- Kane, G.L. (1987). *Modern Elementary Particle Physics*. Perseus Books. ISBN 0-201-11749-5.
- Kleinert, H.; Schulte-Frohlinde, Verena (2001). *Critical Properties of φ^4 -Theories*^[4]. World Scientific. ISBN 981-02-4658-7.
- Kleinert, H. (2008). *Multivalued Fields in Condensed Matter, Electrodynamics, and Gravitation*^[5]. World Scientific. ISBN 978-981-279-170-2.
- Loudon, R (1983). *The Quantum Theory of Light*. Oxford University Press. ISBN 0-19-851155-8.
- Mandl, F.; Shaw, G. (1993). *Quantum Field Theory*. John Wiley & Sons. ISBN 0-0471-94186-7.
- Peskin, M.; Schroeder, D. (1995). *An Introduction to Quantum Field Theory*. Westview Press. ISBN 0-201-50397-2.
- Ryder, L.H. (1985). *Quantum Field Theory*. Cambridge University Press. ISBN 0-521-33859-X.
- Srednicki, Mark (2007) *Quantum Field Theory*.^[6] Cambridge Univ. Press.
- Yndurain, F.J. (1996). *Relativistic Quantum Mechanics and Introduction to Field Theory* (1st ed.). Springer. ISBN 978-3540604532.
- Zee, A. (2003). *Quantum Field Theory in a Nutshell*. Princeton University Press. ISBN ISBN 0-691-01019-6.

Advanced texts:

- Bogoliubov, N.; Logunov, A.A.; Oksak, A.I.; Todorov, I.T. (1990). *General Principles of Quantum Field Theory*. Kluwer Academic Publishers. ISBN 978-0792305408.
- Weinberg, S. (1995). *The Quantum Theory of Fields*. **1–3**. Cambridge University Press.

Articles:

- Gerard 't Hooft (2007) "The Conceptual Basis of Quantum Field Theory^[7]" in Butterfield, J., and John Earman, eds., *Philosophy of Physics, Part A*. Elsevier: 661-730.
- Frank Wilczek (1999) "Quantum field theory,^[8]" *Reviews of Modern Physics* 71: S83-S95. Also doi=10.1103/Rev. Mod. Phys. 71 .

External links

- Stanford Encyclopedia of Philosophy: "Quantum Field Theory,^[9]" by Meinard Kuhlmann.
- Siegel, Warren, 2005. *Fields*.^[10] A free text, also available from arXiv:hep-th/9912205.
- Pedagogic Aids to Quantum Field Theory^[11]. Click on "Introduction" for a simplified introduction suitable for someone familiar with quantum mechanics.
- Free condensed matter books and notes^[12].
- Quantum field theory texts^[13], a list with links to amazon.com.
- Quantum Field Theory^[14] by P. J. Mulders
- Quantum Field Theory^[15] by David Tong
- Quantum Field Theory Video Lectures^[16] by David Tong
- Quantum Field Theory Lecture Notes^[17] by Michael Luke
- Quantum Field Theory Video Lectures^[18] by Sidney R. Coleman

References

- [1] Weinberg, S. Quantum Field Theory, Vols. I to III, 2000, Cambridge University Press: Cambridge, UK.
- [2] Dirac, P.A.M. (1927). *The Quantum Theory of the Emission and Absorption of Radiation*, Proceedings of the Royal Society of London, Series A, Vol. 114, p. 243.
- [3] Abraham Pais, *Inward Bound: Of Matter and Forces in the Physical World* ISBN 0-19-851997-4. Pais recounts how his astonishment at the rapidity with which Feynman could calculate using his method. Feynman's method is now part of the standard methods for physicists.
- [4] <http://users.physik.fu-berlin.de/~kleinert/re.html#B6>
- [5] http://users.physik.fu-berlin.de/~kleinert/public_html/kleiner_reb11/psfiles/mvf.pdf
- [6] <http://www.cambridge.org/us/catalogue/catalogue.asp?isbn=0521864496>
- [7] <http://www.phys.uu.nl/~thoof/lectures/basisqft.pdf>
- [8] <http://arxiv.org/abs/hep-th/9803075>
- [9] <http://plato.stanford.edu/entries/quantum-field-theory/>
- [10] <http://insti.physics.sunysb.edu/%7Esiegel/errata.html>
- [11] <http://quantumfieldtheory.info>
- [12] <http://www.freebookcentre.net/Physics/Condensed-Matter-Books.html>
- [13] <http://motls.blogspot.com/2006/01/qft-didactics.html>
- [14] <http://www.nat.vu.nl/~mulders/QFT-0.pdf>
- [15] <http://damtp.cam.ac.uk/user/tong/qft/qft.pdf>
- [16] http://pirsa.org/index.php?p=speaker&name=David_Tong
- [17] http://www2.physics.utoronto.ca/~luke/PHY2403/References_files/lecturenotes.pdf
- [18] <http://www.physics.harvard.edu/about/Phys253.html>

Algebraic quantum field theory

The **Haag-Kastler** axiomatic framework for quantum field theory, named after Rudolf Haag and Daniel Kastler, is an application to *local* quantum physics of C^* -algebra theory. It is therefore also known as **Algebraic Quantum Field Theory (AQFT)**. The axioms are stated in terms of an algebra given for every open set in Minkowski space, and mappings between those.

Let **Mink** be the category of open subsets of Minkowski space M with inclusion maps as morphisms. We are given a covariant functor $\mathcal{A}$ from **Mink** to $\mathbf{uC}^*\mathbf{alg}$, the category of unital C^* algebras, such that every morphism in **Mink** maps to a monomorphism in $\mathbf{uC}^*\mathbf{alg}$ (*isotony*).

The Poincaré group acts continuously on **Mink**. There exists a pullback of this action, which is continuous in the norm topology of $\mathcal{A}(M)$ (Poincaré covariance).

Minkowski space has a causal structure. If an open set V lies in the causal complement of an open set U , then the image of the maps

$$\mathcal{A}(i_{U,U \cup V})$$

and

$$\mathcal{A}(i_{V,U \cup V})$$

commute (spacelike commutativity). If $\bar{U}$ is the causal completion of an open set U , then $\mathcal{A}(i_{U,\bar{U}})$ is an isomorphism (primitive causality).

A state with respect to a C^* -algebra is a positive linear functional over it with unit norm. If we have a state over $\mathcal{A}(M)$, we can take the "partial trace" to get states associated with $\mathcal{A}(U)$ for each open set via the net monomorphism. It's easy to show the states over the open sets form a presheaf structure.

According to the GNS construction, for each state, we can associate a Hilbert space representation of $\mathcal{A}(M)$. Pure states correspond to irreducible representations and mixed states correspond to reducible representations. Each irreducible (up to equivalence) is called a superselection sector. We assume there is a pure state called the vacuum such that the Hilbert space associated with it is a unitary representation of the Poincaré group compatible with the Poincaré covariance of the net such that if we look at the Poincaré algebra, the spectrum with respect to

energy-momentum (corresponding to spacetime translations) lies on and in the positive light cone. This is the vacuum sector.

Suggested reading

- Haag, Rudolf (1992). *Local Quantum Physics: Fields, Particles, Algebras*. Springer.

External links

- Local Quantum Physics Crossroads ^[1] - A network of scientists working on Local Quantum Physics
- Algebraic Quantum Field Theory ^[2] - AQFT resources at the University of Hamburg

References

[1] <http://www.lqp.uni-goettingen.de/>

[2] <http://unith.desy.de/research/aqft/>

Local quantum field theory

The **Haag-Kastler** axiomatic framework for quantum field theory, named after Rudolf Haag and Daniel Kastler, is an application to *local* quantum physics of C^* -algebra theory. It is therefore also known as **Algebraic Quantum Field Theory (AQFT)**. The axioms are stated in terms of an algebra given for every open set in Minkowski space, and mappings between those.

Let **Mink** be the category of open subsets of Minkowski space M with inclusion maps as morphisms. We are given a covariant functor $\mathcal{A}$ from **Mink** to **uC*alg**, the category of unital C^* algebras, such that every morphism in **Mink** maps to a monomorphism in **uC*alg** (*isotony*).

The Poincaré group acts continuously on **Mink**. There exists a pullback of this action, which is continuous in the norm topology of $\mathcal{A}(M)$ (Poincaré covariance).

Minkowski space has a causal structure. If an open set V lies in the causal complement of an open set U , then the image of the maps

$$\mathcal{A}(i_{U,U \cup V})$$

and

$$\mathcal{A}(i_{V,U \cup V})$$

commute (spacelike commutativity). If $\bar{U}$ is the causal completion of an open set U , then $\mathcal{A}(i_{U,\bar{U}})$ is an isomorphism (primitive causality).

A state with respect to a C^* -algebra is a positive linear functional over it with unit norm. If we have a state over $\mathcal{A}(M)$, we can take the "partial trace" to get states associated with $\mathcal{A}(U)$ for each open set via the net monomorphism. It's easy to show the states over the open sets form a presheaf structure.

According to the GNS construction, for each state, we can associate a Hilbert space representation of $\mathcal{A}(M)$. Pure states correspond to irreducible representations and mixed states correspond to reducible representations. Each irreducible (up to equivalence) is called a superselection sector. We assume there is a pure state called the vacuum such that the Hilbert space associated with it is a unitary representation of the Poincaré group compatible with the Poincaré covariance of the net such that if we look at the Poincaré algebra, the spectrum with respect to energy-momentum (corresponding to spacetime translations) lies on and in the positive light cone. This is the vacuum sector.

Suggested reading

- Haag, Rudolf (1992). *Local Quantum Physics: Fields, Particles, Algebras*. Springer.

External links

- Local Quantum Physics Crossroads ^[1] - A network of scientists working on Local Quantum Physics
- Algebraic Quantum Field Theory ^[2] - AQFT resources at the University of Hamburg

Algebraic logic

In mathematical logic, **algebraic logic** is the study of logic presented in an algebraic style.

Algebras as models of logics

Algebraic logic treats algebraic structures, often bounded lattices, as models (interpretations) of certain logics, making logic a branch of order theory.

In algebraic logic:

- Variables are tacitly universally quantified over some universe of discourse. There are no existentially quantified variables or open formulas;
- Terms are built up from variables using primitive and defined operations. There are no connectives;
- Formulas, built from terms in the usual way, can be equated if they are logically equivalent. To express a tautology, equate a formula with a truth value;
- The rules of proof are the substitution of equals for equals, and uniform replacement. Modus ponens remains valid, but is seldom employed.

In the table below, the left column contains one or more logical or mathematical systems, and the algebraic structure which are its models are shown on the right in the same row. Some of these structures are either Boolean algebras or proper extensions thereof. Modal and other nonclassical logics are typically modeled by what are called "Boolean algebras with operators."

Algebraic formalisms going beyond first-order logic in at least some respects include:

- Combinatory logic, having the expressive power of set theory;
- Relation algebra, arguably the paradigmatic algebraic logic, can express Peano arithmetic and most axiomatic set theories, including the canonical ZFC.

logical system	its models
Classical sentential logic	Lindenbaum-Tarski algebra Two-element Boolean algebra
Intuitionistic propositional logic	Heyting algebra
Łukasiewicz logic	MV-algebra
Modal logic K	Modal algebra
Lewis's S4	Interior algebra
Lewis's S5; Monadic predicate logic	Monadic Boolean algebra
First-order logic	Cylindric algebra Polyadic algebra Predicate functor logic
Set theory	Combinatory logic Relation algebra

History

On the history of algebraic logic before World War II, see Brady (2000) and Grattan-Guinness (2000) and their ample references. On the postwar history, see Maddux (1991) and Quine (1976).

Algebraic logic has at least two meanings:

- The study of Boolean algebra, begun by George Boole, and of relation algebra, begun by Augustus DeMorgan, extended by Charles Sanders Peirce, and taking definitive form in the work of Ernst Schröder;
- Abstract algebraic logic, a branch of contemporary mathematical logic.

Perhaps surprisingly, algebraic logic is the oldest approach to formal logic, arguably beginning with a number of memoranda Leibniz wrote in the 1680s, some of which were published in the 19th century and translated into English by Clarence Lewis in 1918. But nearly all of Leibniz's known work on algebraic logic was published only in 1903, after Louis Couturat discovered it in Leibniz's Nachlass. Parkinson (1966) and Loemker (1969) translated selections from Couturat's volume into English.

Brady (2000) discusses the rich historical connections between algebraic logic and model theory. The founders of model theory, Ernst Schröder and Leopold Löwenheim, were logicians in the algebraic tradition. Alfred Tarski, the founder of set theoretic model theory as a major branch of contemporary mathematical logic, also:

- Co-discovered Lindenbaum-Tarski algebra;
- Invented cylindric algebra;
- Wrote the 1940 paper that revived relation algebra, and that can be seen as the starting point of abstract algebraic logic.

Modern mathematical logic began in 1847, with two pamphlets whose respective authors were Augustus DeMorgan and George Boole. They, and later C.S. Peirce, Hugh MacColl, Frege, Peano, Bertrand Russell, and A. N. Whitehead all shared Leibniz's dream of combining symbolic logic, mathematics, and philosophy. Relation algebra is arguably the culmination of Leibniz's approach to logic. With the exception of some writings by Leopold Löwenheim and Thoralf Skolem, algebraic logic went into eclipse soon after the 1910-13 publication of *Principia Mathematica*, not to revive until Tarski's 1940 reexposition of relation algebra.

Leibniz had no influence on the rise of algebraic logic because his logical writings were little studied before the Parkinson and Loemker translations. Our present understanding of Leibniz the logician stems mainly from the work of Wolfgang Lenzen, summarized in Lenzen (2004).^[1] To see how present-day work in logic and metaphysics can draw inspiration from, and shed light on, Leibniz's thought, see Zalta (2000).^[2]

See also

- Abstract algebraic logic
- Algebraic structure
- Boolean algebra (logic)
- Boolean algebra (structure)
- Cylindric algebra
- Lindenbaum-Tarski algebra
- Mathematical logic
- Model theory
- Monadic Boolean algebra
- Predicate functor logic
- Relation algebra
- Universal algebra

References

- Brady, Geraldine, 2000. *From Peirce to Skolem: A neglected chapter in the history of logic*. North-Holland/Elsevier Science BV: catalog page ^[3], Amsterdam, Netherlands, 625 pages.
- Burris, Stanley, 2009. The Algebra of Logic Tradition ^[4]. Stanford Encyclopedia of Philosophy.
- Ivor Grattan-Guinness, 2000. *The Search for Mathematical Roots*. Princeton Univ. Press.
- Lenzen, Wolfgang, 2004, "Leibniz's Logic ^[1]" in Gabbay, D., and Woods, J., eds., *Handbook of the History of Logic, Vol. 3: The Rise of Modern Logic from Leibniz to Frege*. North-Holland: 1-84.
- Loemker, Leroy (1969 (1956)), *Leibniz: Philosophical Papers and Letters*, Reidel.
- Roger Maddux, 1991, "The Origin of Relation Algebras in the Development and Axiomatization of the Calculus of Relations," *Studia Logica* 50: 421-55.
- Parkinson, G.H.R., 1966. *Leibniz: Logical Papers*. Oxford Uni. Press.
- Willard Quine, 1976, "Algebraic Logic and Predicate Functors" in *The Ways of Paradox*. Harvard Univ. Press: 283-307.
- Zalta, E. N., 2000, "A (Leibnizian) Theory of Concepts ^[5]," *Philosophiegeschichte und logische Analyse / Logical Analysis and History of Philosophy* 3: 137-183.

External links

- Stanford Encyclopedia of Philosophy: "Propositional Consequence Relations and Algebraic Logic ^[6]" -- by Ramon Jansana.

References

- [1] <http://www.philosophie.uni-osnabrueck.de/Publikationen%20Lenzen/Lenzen%20Leibniz%20Logic.pdf>
- [2] <http://mally.stanford.edu/Papers/leibniz.pdf>
- [3] http://www.elsevier.com/wps/find/bookdescription.cws_home/621535/description
- [4] <http://plato.stanford.edu/entries/algebra-logic-tradition/>
- [5] <http://mally.stanford.edu/leibniz.pdf>
- [6] <http://plato.stanford.edu/entries/consequence-algebraic/>

Quantum logic

In quantum mechanics, **quantum logic** is a set of rules for reasoning about propositions which takes the principles of quantum theory into account. This research area and its name originated in the 1936 paper by Garrett Birkhoff and John von Neumann, who were attempting to reconcile the apparent inconsistency of classical boolean logic with the facts concerning the measurement of complementary variables in quantum mechanics, such as position and momentum.

Quantum logic can be formulated either as a modified version of propositional logic or as a noncommutative and non-associative many-valued (MV) logic^{[1] [2] [3] [4] [5]}.

Quantum logic has some properties which clearly distinguish it from classical logic, most notably, the failure of the distributive law of propositional logic:

$$p \text{ and } (q \text{ or } r) = (p \text{ and } q) \text{ or } (p \text{ and } r),$$

where the symbols p , q and r are propositional variables. To illustrate why the distributive law fails, consider a particle moving on a line and let

$$p = \text{"the particle is moving to the right"}$$

$$q = \text{"the particle is in the interval } [-1,1]\text{"}$$

$$r = \text{"the particle is not in the interval } [-1,1]\text{"}$$

then the proposition " q or r " is true, so

$$p \text{ and } (q \text{ or } r) = p$$

On the other hand, the propositions " p and q " and " p and r " are both false, since they assert tighter restrictions on simultaneous values of position and momentum than is allowed by the uncertainty principle. So,

$$(p \text{ and } q) \text{ or } (p \text{ and } r) = \text{false}$$

Thus the distributive law fails.

Quantum logic has been proposed as the correct logic for propositional inference generally, most notably by the philosopher Hilary Putnam, at least at one point in his career. This thesis was an important ingredient in Putnam's paper *Is Logic Empirical?* in which he analysed the epistemological status of the rules of propositional logic. Putnam attributes the idea that anomalies associated to quantum measurements originate with anomalies in the logic of physics itself to the physicist David Finkelstein. However, this idea had been around for some time and had been revived several years earlier by George Mackey's work on group representations and symmetry.

The more common view regarding quantum logic, however, is that it provides a formalism for relating observables, system preparation filters and states. In this view, the quantum logic approach resembles more closely the C*-algebraic approach to quantum mechanics; in fact with some minor technical assumptions it can be subsumed by it. The similarities of the quantum logic formalism to a system of deductive logic may then be regarded more as a curiosity than as a fact of fundamental philosophical importance. A more modern approach to the structure of quantum logic is to assume that it is a diagram – in the sense of category theory – of classical logics (see David Edwards).

Introduction

In his classic treatise *Mathematical Foundations of Quantum Mechanics*, John von Neumann noted that projections on a Hilbert space can be viewed as propositions about physical observables. The set of principles for manipulating these quantum propositions was called *quantum logic* by von Neumann and Birkhoff. In his book (also called *Mathematical Foundations of Quantum Mechanics*) G. Mackey attempted to provide a set of axioms for this propositional system as an orthocomplemented lattice. Mackey viewed elements of this set as potential *yes or no*

questions an observer might ask about the state of a physical system, questions that would be settled by some measurement. Moreover Mackey defined a physical observable in terms of these basic questions. Mackey's axiom system is somewhat unsatisfactory though, since it assumes that the partially ordered set is actually given as the orthocomplemented closed subspace lattice of a separable Hilbert space. Piron, Ludwig and others have attempted to give axiomatizations which do not require such explicit relations to the lattice of subspaces.

The remainder of this article assumes the reader is familiar with the spectral theory of self-adjoint operators on a Hilbert space. However, the main ideas can be understood using the finite-dimensional spectral theorem.

Projections as propositions

The so-called *Hamiltonian* formulations of classical mechanics have three ingredients: *states*, *observables* and *dynamics*. In the simplest case of a single particle moving in $\mathbf{R}^3$, the state space is the position-momentum space $\mathbf{R}^6$. We will merely note here that an observable is some real-valued function f on the state space. Examples of observables are position, momentum or energy of a particle. For classical systems, the value $f(x)$, that is the value of f for some particular system state x , is obtained by a process of measurement of f . The propositions concerning a classical system are generated from basic statements of the form

- Measurement of f yields a value in the interval $[a, b]$ for some real numbers a, b .

It follows easily from this characterization of propositions in classical systems that the corresponding logic is identical to that of some Boolean algebra of subsets of the state space. By logic in this context we mean the rules that relate set operations and ordering relations, such as de Morgan's laws. These are analogous to the rules relating boolean conjunctives and material implication in classical propositional logic. For technical reasons, we will also assume that the algebra of subsets of the state space is that of all Borel sets. The set of propositions is ordered by the natural ordering of sets and has a complementation operation. In terms of observables, the complement of the proposition $\{f \geq a\}$ is $\{f < a\}$.

We summarize these remarks as follows:

- The proposition system of a classical system is a lattice with a distinguished *orthocomplementation* operation: The lattice operations of *meet* and *join* are respectively set intersection and set union. The orthocomplementation operation is set complement. Moreover this lattice is *sequentially complete*, in the sense that any sequence $\{E_i\}_i$ of elements of the lattice has a least upper bound, specifically the set-theoretic union:

$$\text{LUB}(\{E_i\}) = \bigcup_{i=1}^{\infty} E_i.$$

In the Hilbert space formulation of quantum mechanics as presented by von Neumann, a physical observable is represented by some (possibly unbounded) densely-defined self-adjoint operator A on a Hilbert space H . A has a spectral decomposition, which is a projection-valued measure E defined on the Borel subsets of $\mathbf{R}$. In particular, for any bounded Borel function f , the following equation holds:

$$f(A) = \int_{\mathbf{R}} f(\lambda) dE(\lambda).$$

In case f is the indicator function of an interval $[a, b]$, the operator $f(A)$ is a self-adjoint projection, and can be interpreted as the quantum analogue of the classical proposition

- Measurement of A yields a value in the interval $[a, b]$.

The propositional lattice of a quantum mechanical system

This suggests the following quantum mechanical replacement for the orthocomplemented lattice of propositions in classical mechanics. This is essentially Mackey's *Axiom VII*:

- The orthocomplemented lattice Q of propositions of a quantum mechanical system is the lattice of closed subspaces of a complex Hilbert space H where orthocomplementation of V is the orthogonal complement $V^\perp$.

Q is also sequentially complete: any pairwise disjoint sequence $\{V_i\}_i$ of elements of Q has a least upper bound. Here disjointness of W_1 and W_2 means W_2 is a subspace of $W_1^\perp$. The least upper bound of $\{V_i\}_i$ is the closed internal direct sum.

Henceforth we identify elements of Q with self-adjoint projections on the Hilbert space H .

The structure of Q immediately points to a difference with the partial order structure of a classical proposition system. In the classical case, given a proposition p , the equations

$$I = p \vee q$$

$$0 = p \wedge q$$

have exactly one solution, namely the set-theoretic complement of p . In these equations I refers to the atomic proposition which is identically true and 0 the atomic proposition which is identically false. In the case of the lattice of projections there are infinitely many solutions to the above equations.

Having made these preliminary remarks, we turn everything around and attempt to define observables within the projection lattice framework and using this definition establish the correspondence between self-adjoint operators and observables: A *Mackey observable* is a countably additive homomorphism from the orthocomplemented lattice of the Borel subsets of $\mathbf{R}$ to Q . To say the mapping φ is a countably additive homomorphism means that for any sequence $\{S_i\}_i$ of pairwise disjoint Borel subsets of $\mathbf{R}$, $\{\varphi(S_i)\}_i$ are pairwise orthogonal projections and

$$\varphi\left(\bigcup_{i=1}^{\infty} S_i\right) = \sum_{i=1}^{\infty} \varphi(S_i).$$

Theorem. There is a bijective correspondence between Mackey observables and densely-defined self-adjoint operators on H .

This is the content of the spectral theorem as stated in terms of spectral measures.

Statistical structure

Imagine a forensics lab which has some apparatus to measure the speed of a bullet fired from a gun. Under carefully controlled conditions of temperature, humidity, pressure and so on the same gun is fired repeatedly and speed measurements taken. This produces some distribution of speeds. Though we will not get exactly the same value for each individual measurement, for each cluster of measurements, we would expect the experiment to lead to the same distribution of speeds. In particular, we can expect to assign probability distributions to propositions such as $\{a \leq \text{speed} \leq b\}$. This leads naturally to propose that under controlled conditions of preparation, the measurement of a classical system can be described by a probability measure on the state space. This same statistical structure is also present in quantum mechanics.

A *quantum probability measure* is a function P defined on Q with values in $[0,1]$ such that $P(0)=0$, $P(I)=1$ and if $\{E_i\}_i$ is a sequence of pairwise orthogonal elements of Q then

$$P\left(\sum_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

The following highly non-trivial theorem is due to Andrew Gleason:

Theorem. Suppose H is a separable Hilbert space of complex dimension at least 3. Then for any quantum probability measure on Q there exists a unique trace class operator S such that

$$P(E) = \text{Tr}(SE)$$

for any self-adjoint projection E .

The operator S is necessarily non-negative (that is all eigenvalues are non-negative) and of trace 1. Such an operator is often called a *density operator*.

Physicists commonly regard a density operator as being represented by a (possibly infinite) density matrix relative to some orthonormal basis.

For more information on statistics of quantum systems, see quantum statistical mechanics.

Automorphisms

An *automorphism* of $\mathcal{Q}$ is a bijective mapping $\alpha: \mathcal{Q} \rightarrow \mathcal{Q}$ which preserves the orthocomplemented structure of $\mathcal{Q}$, that is

$$\alpha\left(\sum_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \alpha(E_i)$$

for any sequence $\{E_i\}_i$ of pairwise orthogonal self-adjoint projections. Note that this property implies monotonicity of α . If P is a quantum probability measure on $\mathcal{Q}$, then $E \rightarrow \alpha(E)$ is also a quantum probability measure on $\mathcal{Q}$. By the Gleason theorem characterizing quantum probability measures quoted above, any automorphism α induces a mapping α^* on the density operators by the following formula:

$$\text{Tr}(\alpha^*(S)E) = \text{Tr}(S\alpha(E)).$$

The mapping α^* is bijective and preserves convex combinations of density operators. This means

$$\alpha^*(r_1 S_1 + r_2 S_2) = r_1 \alpha^*(S_1) + r_2 \alpha^*(S_2)$$

whenever $1 = r_1 + r_2$ and r_1, r_2 are non-negative real numbers. Now we use a theorem of Richard V. Kadison:

Theorem. Suppose β is a bijective map from density operators to density operators which is convexity preserving. Then there is an operator U on the Hilbert space which is either linear or conjugate-linear, preserves the inner product and is such that

$$\beta(S) = USU^*$$

for every density operator S . In the first case we say U is unitary, in the second case U is anti-unitary.

Remark. This note is included for technical accuracy only, and should not concern most readers. The result quoted above is not directly stated in Kadison's paper, but can be reduced to it by noting first that β extends to a positive trace preserving map on the trace class operators, then applying duality and finally applying a result of Kadison's paper.

The operator U is not quite unique; if r is a complex scalar of modulus 1, then rU will be unitary or anti-unitary if U is and will implement the same automorphism. In fact, this is the only ambiguity possible.

It follows that automorphisms of $\mathcal{Q}$ are in bijective correspondence to unitary or anti-unitary operators modulo multiplication by scalars of modulus 1. Moreover, we can regard automorphisms in two equivalent ways: as operating on states (represented as density operators) or as operating on $\mathcal{Q}$.

Non-relativistic dynamics

In non-relativistic physical systems, there is no ambiguity in referring to time evolution since there is a global time parameter. Moreover an isolated quantum system evolves in a deterministic way: if the system is in a state S at time t then at time $s > t$, the system is in a state $F_{s,t}(S)$. Moreover, we assume

- The dependence is reversible: The operators $F_{s,t}$ are bijective.
- The dependence is homogeneous: $F_{s,t} = F_{s-t,0}$.
- The dependence is convexity preserving: That is, each $F_{s,t}(S)$ is convexity preserving.
- The dependence is weakly continuous: The mapping $\mathbf{R} \rightarrow \mathbf{R}$ given by $t \rightarrow \text{Tr}(F_{s,t}(S) E)$ is continuous for every E in $\mathcal{Q}$.

By Kadison's theorem, there is a 1-parameter family of unitary or anti-unitary operators $\{U_t\}_t$ such that

$$F_{s,t}(S) = U_{s-t} S U_{s-t}^*$$

In fact,

Theorem. Under the above assumptions, there is a strongly continuous 1-parameter group of unitary operators $\{U_t\}_t$ such that the above equation holds.

Note that it easily from uniqueness from Kadison's theorem that

$$U_{t+s} = \sigma(t, s) U_t U_s$$

where $\sigma(t, s)$ has modulus 1. Now the square of an anti-unitary is a unitary, so that all the U_t are unitary. The remainder of the argument shows that $\sigma(t, s)$ can be chosen to be 1 (by modifying each U_t by a scalar of modulus 1.)

Pure states

A convex combinations of statistical states S_1 and S_2 is a state of the form $S = p_1 S_1 + p_2 S_2$ where p_1, p_2 are non-negative and $p_1 + p_2 = 1$. Considering the statistical state of system as specified by lab conditions used for its preparation, the convex combination S can be regarded as the state formed in the following way: toss a biased coin with outcome probabilities p_1, p_2 and depending on outcome choose system prepared to S_1 or S_2 .

Density operators form a convex set. The convex set of density operators has extreme points; these are the density operators given by a projection onto a one-dimensional space. To see that any extreme point is such a projection, note that by the spectral theorem S can be represented by a diagonal matrix; since S is non-negative all the entries are non-negative and since S has trace 1, the diagonal entries must add up to 1. Now if it happens that the diagonal matrix has more than one non-zero entry it is clear that we can express it as a convex combination of other density operators.

The extreme points of the set of density operators are called pure states. If S is the projection on the 1-dimensional space generated by a vector ψ of norm 1 then

$$\text{Tr}(SE) = \langle E\psi | \psi \rangle$$

for any E in $\mathcal{Q}$. In physics jargon, if

$$S = |\psi\rangle\langle\psi|,$$

where ψ has norm 1, then

$$\text{Tr}(SE) = \langle \psi | E | \psi \rangle.$$

Thus pure states can be identified with *rays* in the Hilbert space H .

The measurement process

Consider a quantum mechanical system with lattice Q which is in some statistical state given by a density operator S . This essentially means an ensemble of systems specified by a repeatable lab preparation process. The result of a cluster of measurements intended to determine the truth value of proposition E , is just as in the classical case, a probability distribution of truth values **T** and **F**. Say the probabilities are p for **T** and $q = 1 - p$ for **F**. By the previous section $p = \text{Tr}(SE)$ and $q = \text{Tr}(S(I - E))$.

Perhaps the most fundamental difference between classical and quantum systems is the following: regardless of what process is used to determine E immediately after the measurement the system will be in one of two statistical states:

- If the result of the measurement is **T**

$$\frac{1}{\text{Tr}(ES)} ESE.$$

- If the result of the measurement is **F**

$$\frac{1}{\text{Tr}((I - E)S)} (I - E)S(I - E).$$

(We leave to the reader the handling of the degenerate cases in which the denominators may be 0.) We now form the convex combination of these two ensembles using the relative frequencies p and q . We thus obtain the result that the measurement process applied to a statistical ensemble in state S yields another ensemble in statistical state:

$$M_E(S) = ESE + (I - E)S(I - E).$$

We see that a pure ensemble becomes a mixed ensemble after measurement. Measurement, as described above, is a special case of quantum operations.

Limitations

Quantum logic derived from propositional logic provides a satisfactory foundation for a theory of reversible quantum processes. Examples of such processes are the covariance transformations relating two frames of reference, such as change of time parameter or the transformations of special relativity. Quantum logic also provides a satisfactory understanding of density matrices. Quantum logic can be stretched to account for some kinds of measurement processes corresponding to answering yes-no questions about the state of a quantum system. However, for more general kinds of measurement operations (that is quantum operations), a more complete theory of filtering processes is necessary. Such an approach is provided by the consistent histories formalism. On the other hand, quantum logics derived from MV-logic extend its range of applicability to irreversible quantum processes and/or 'open' quantum systems.

In any case, these quantum logic formalisms must be generalized in order to deal with super-geometry (which is needed to handle Fermi-fields) and non-commutative geometry (which is needed in string theory and quantum gravity theory). Both of these theories use a partial algebra with an "integral" or "trace". The elements of the partial algebra are not observables; instead the "trace" yields "greens functions" which generate scattering amplitudes. One thus obtains a local S-matrix theory (see D. Edwards).

Since around 1978 the Flato school (see F. Bayen) has been developing an alternative to the quantum logics approach called deformation quantization (see Weyl quantization).

In 2004, Prakash Panangaden described how to capture the kinematics of quantum causal evolution using System BV, a deep inference logic originally developed for use in structural proof theory.^[6] Alessio Guglielmi, Lutz Straßburger, and Richard Blute have also done work in this area.^[7]

See also

- Mathematical formulation of quantum mechanics
- Multi-valued logic
- Quasi-set theory
- HPO formalism (An approach to temporal quantum logic)
- Quantum field theory

Further reading

- S. Auyang, *How is Quantum Field Theory Possible?*, Oxford University Press, 1995.
- F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz and D. Sternheimer, *Deformation theory and quantization I,II*, Ann. Phys. (N.Y.), 111 (1978) pp. 61–110, 111–151.
- G. Birkhoff and J. von Neumann, *The Logic of Quantum Mechanics*, Annals of Mathematics, Vol. 37, pp. 823–843, 1936.
- D. Cohen, *An Introduction to Hilbert Space and Quantum Logic*, Springer-Verlag, 1989. This is a thorough but elementary and well-illustrated introduction, suitable for advanced undergraduates.
- David Edwards, *The Mathematical Foundations of Quantum Mechanics*, Synthese, Volume 42, Number 1/September, 1979, pp. 1–70.
- D. Edwards, *The Mathematical Foundations of Quantum Field Theory: Fermions, Gauge Fields, and Super-symmetry, Part I: Lattice Field Theories*, International J. of Theor. Phys., Vol. 20, No. 7 (1981).
- D. Finkelstein, *Matter, Space and Logic*, Boston Studies in the Philosophy of Science Vol. V, 1969
- A. Gleason, *Measures on the Closed Subspaces of a Hilbert Space*, Journal of Mathematics and Mechanics, 1957.
- R. Kadison, *Isometries of Operator Algebras*, Annals of Mathematics, Vol. 54, pp. 325–338, 1951
- G. Ludwig, *Foundations of Quantum Mechanics*, Springer-Verlag, 1983.
- G. Mackey, *Mathematical Foundations of Quantum Mechanics*, W. A. Benjamin, 1963 (paperback reprint by Dover 2004).
- J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955. Reprinted in paperback form.
- R. Omnès, *Understanding Quantum Mechanics*, Princeton University Press, 1999. An extraordinarily lucid discussion of some logical and philosophical issues of quantum mechanics, with careful attention to the history of the subject. Also discusses consistent histories.
- N. Papanikolaou, *Reasoning Formally About Quantum Systems: An Overview*, ACM SIGACT News, 36(3), pp. 51–66, 2005.
- C. Piron, *Foundations of Quantum Physics*, W. A. Benjamin, 1976.
- H. Putnam, *Is Logic Empirical?*, Boston Studies in the Philosophy of Science Vol. V, 1969
- H. Weyl, *The Theory of Groups and Quantum Mechanics*, Dover Publications, 1950.

External links

- Stanford Encyclopedia of Philosophy entry on Quantum Logic and Probability Theory ^[8]

References

- [1] http://arxiv.org/PS_cache/quant-ph/pdf/0101/0101028v2.pdf Maria Luisa Dalla Chiara and Roberto Giuntini. 2008. *Quantum Logic.*, 102 pages PDF
- [2] Dalla Chiara, M. L. and Giuntini, R.: 1994, Unsharp quantum logics, *Foundations of Physics.*, **24**, 1161–1177.
- [3] <http://planetphysics.org/encyclopedia/QuantumLMAlgebraicLogic.html> I. C. Baianu. 2009. Quantum LMn Algebraic Logic.
- [4] Georgescu, G. and C. Vraciu. 1970, On the characterization of centered Łukasiewicz algebras., *J. Algebra*, **16**: 486-495.
- [5] Georgescu, G. 2006, N-valued Logics and Łukasiewicz-Moisil Algebras, *Axiomathes*, **16** (1-2): 123-
- [6] <http://cs.bath.ac.uk/ag/p/BVQuantCausEvol.pdf>
- [7] <http://alessio.guglielmi.name/res/cos/crt.html#CQE>
- [8] <http://plato.stanford.edu/entries/qt-quantlog/>

Quantum computer

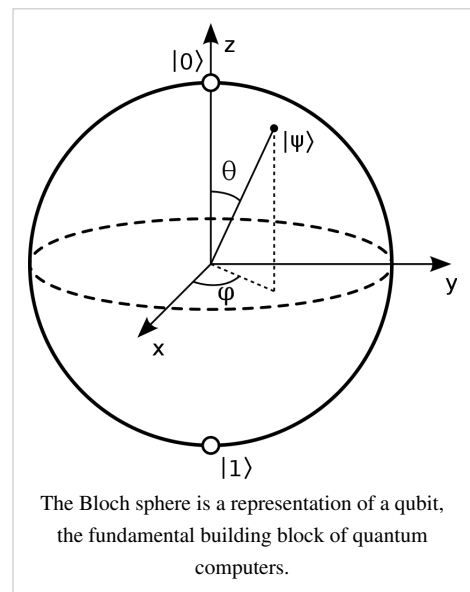
A **quantum computer** is a device for computation that makes direct use of quantum mechanical phenomena, such as superposition and entanglement, to perform operations on data. Quantum computers are different from traditional computers based on transistors. The basic principle behind quantum computation is that quantum properties can be used to represent data and perform operations on these data.^[1] A theoretical model is the quantum Turing machine, also known as the universal quantum computer.

Although quantum computing is still in its infancy, experiments have been carried out in which quantum computational operations were executed on a very small number of qubits (**quantum bit**). Both practical and theoretical research continues, and many national government and military funding agencies support quantum computing research to develop quantum computers for both civilian and national security purposes, such as cryptanalysis.^[2]

If large-scale quantum computers can be built, they will be able to solve certain problems much faster than any current classical computers (for example Shor's algorithm). Quantum computers however do not allow one to compute functions that are not theoretically computable by classical computers, i.e. they do not alter the Church–Turing thesis. The gain is only in efficiency.

Basis

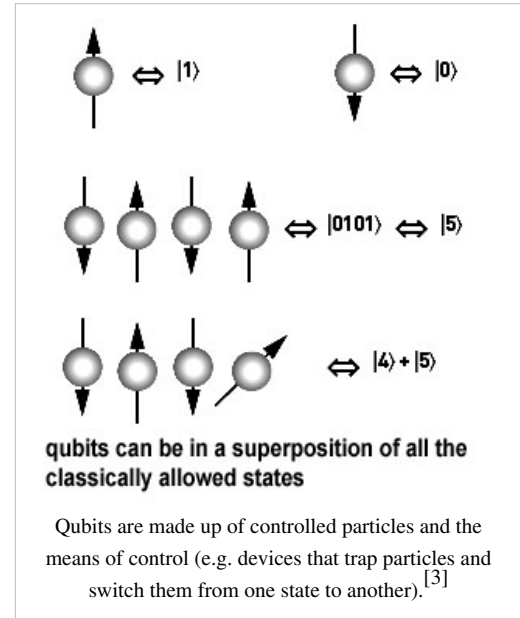
A classical computer has a memory made up of bits, where each bit represents either a one or a zero. A quantum computer maintains a sequence of qubits. A single qubit can represent a one, a zero, or, crucially, any quantum superposition of these; moreover, a pair of qubits can be in any quantum superposition of 4 states, and three qubits in any superposition of 8. In general a quantum computer with n qubits can be in an arbitrary superposition of up to 2^n different states simultaneously (this compares to a normal computer that can only be in *one* of these 2^n states at any one time). A quantum computer operates by manipulating those qubits with a fixed sequence of quantum logic gates. The sequence of gates to be applied is called a quantum algorithm.



An example of an implementation of qubits for a quantum computer could start with the use of particles with two spin states: "down" and "up" (typically written $|\downarrow\rangle$ and $|\uparrow\rangle$, or $|0\rangle$ and $|1\rangle$). But in fact any system possessing an observable quantity A which is *conserved* under time evolution and such that A has at least two discrete and sufficiently spaced consecutive eigenvalues, is a suitable candidate for implementing a qubit. This is true because any such system can be mapped onto an effective spin-1/2 system.

Bits vs. qubits

Consider first a classical computer that operates on a three-bit register. The state of the computer at any time is a probability distribution over the $2^3 = 8$ different three-bit strings 000, 001, 010, 011, 100, 101, 110, 111. If it is a deterministic computer, then it is in exactly one of these states with probability 1. However, if it is a probabilistic computer, then there is a possibility of it being in any *one* of a number of different states. We can describe this probabilistic state by eight nonnegative numbers a, b, c, d, e, f, g, h (where a = probability computer is in state 000, b = probability computer is in state 001, etc.). There is a restriction that these probabilities sum to 1.



The state of a three-qubit quantum computer is similarly described by an eight-dimensional vector (a, b, c, d, e, f, g, h) , called a ket. However, instead of adding to one, the sum of the *squares* of the coefficient magnitudes, $|a|^2 + |b|^2 + \dots + |h|^2$, must equal one. Moreover, the coefficients are complex numbers. Since states are represented by complex wavefunctions, two states being added together will undergo interference. This is a key difference between quantum computing and probabilistic classical computing.^[4]

If you measure the three qubits, then you will observe a three-bit string. The probability of measuring a string will equal the squared magnitude of that string's coefficients (using our example, probability that we read state as 000 = $|a|^2$, probability that we read state as 001 = $|b|^2$, etc.). Thus a measurement of the quantum state with coefficients $(a, b, \dots, h)$ gives the classical probability distribution $(|a|^2, |b|^2, \dots, |h|^2)$. We say that the quantum state "collapses" to a classical state.

Note that an eight-dimensional vector can be specified in many different ways, depending on what basis you choose for the space. The basis of three-bit strings 000, 001, ..., 111 is known as the computational basis, and is often convenient, but other bases of unit-length, orthogonal vectors can also be used. Ket notation is often used to make explicit the choice of basis. For example, the state (a, b, c, d, e, f, g, h) in the computational basis can be written as

$$a|000\rangle + b|001\rangle + c|010\rangle + d|011\rangle + e|100\rangle + f|101\rangle + g|110\rangle + h|111\rangle, \quad \text{where, e.g., } |010\rangle = (0, 0, 1, 0, 0, 0, 0, 0).$$

The computational basis for a single qubit (two dimensions) is $|0\rangle = (1, 0)$, $|1\rangle = (0, 1)$, but another common basis are the eigenvectors of the Pauli-x operator: $|+\rangle = \frac{1}{\sqrt{2}}(1, 1)$ and $|-\rangle = \frac{1}{\sqrt{2}}(1, -1)$.

Note that although recording a classical state of n bits, a 2^n -dimensional probability distribution, requires an exponential number of real numbers, practically we can always think of the system as being exactly one of the n -bit strings—we just don't know which one. Quantum mechanically, this is not the case, and all 2^n complex coefficients

need to be kept track of to see how the quantum system evolves. For example, a 300-qubit quantum computer has a state described by 2^{300} (approximately 10^{90}) complex numbers, more than the number of atoms in the observable universe.

Operation

While a classical three-bit state and a quantum three-qubit state are both eight-dimensional vectors, they are manipulated quite differently for classical or quantum computation. For computing in either case, the system must be initialized, for example into the all-zeros string, $|000\rangle$, corresponding to the vector $(1,0,0,0,0,0,0,0)$. In classical randomized computation, the system evolves according to the application of stochastic matrices, which preserve that the probabilities add up to one (i.e., preserve the L1 norm). In quantum computation, on the other hand, allowed operations are unitary matrices, which are effectively rotations (they preserve that the sum of the squares add up to one, the Euclidean or L2 norm). (Exactly what unitaries can be applied depend on the physics of the quantum device.) Consequently, since rotations can be undone by rotating backward, quantum computations are reversible. (Technically, quantum operations can be probabilistic combinations of unitaries, so quantum computation really does generalize classical computation. See quantum circuit for a more precise formulation.)

Finally, upon termination of the algorithm, the result needs to be read off. In the case of a classical computer, we *sample* from the probability distribution on the three-bit register to obtain one definite three-bit string, say 000. Quantum mechanically, we *measure* the three-qubit state, which is equivalent to collapsing the quantum state down to a classical distribution (with the coefficients in the classical state being the squared magnitudes of the coefficients for the quantum state, as described above) followed by sampling from that distribution. Note that this destroys the original quantum state. Many algorithms will only give the correct answer with a certain probability, however by repeatedly initializing, running and measuring the quantum computer, the probability of getting the correct answer can be increased.

For more details on the sequences of operations used for various quantum algorithms, see universal quantum computer, Shor's algorithm, Grover's algorithm, Deutsch-Jozsa algorithm, amplitude amplification, quantum Fourier transform, quantum gate, quantum adiabatic algorithm and quantum error correction.

Potential

Integer factorization is believed to be computationally infeasible with an ordinary computer for large integers that are the product of only a few prime numbers (e.g., products of two 300-digit primes).^[5] By comparison, a quantum computer could efficiently solve this problem using Shor's algorithm to find its factors. This ability would allow a quantum computer to decrypt many of the cryptographic systems in use today, in the sense that there would be a polynomial time (in the number of digits of the integer) algorithm for solving the problem. In particular, most of the popular public key ciphers are based on the difficulty of factoring integers (or the related discrete logarithm problem which can also be solved by Shor's algorithm), including forms of RSA. These are used to protect secure Web pages, encrypted email, and many other types of data. Breaking these would have significant ramifications for electronic privacy and security.

However, other existing cryptographic algorithms don't appear to be broken by these algorithms.^{[6] [7]} Some public-key algorithms are based on problems other than the integer factorization and discrete logarithm problems to which Shor's algorithm applies, like the McEliece cryptosystem based on a problem in coding theory.^{[6] [8]} Lattice based cryptosystems are also not known to be broken by quantum computers, and finding a polynomial time algorithm for solving the dihedral hidden subgroup problem, which would break many lattice based cryptosystems, is a well-studied open problem.^[9] It has been proven that applying Grover's algorithm to break a symmetric (secret key) algorithm by brute force requires roughly $2^{n/2}$ invocations of the underlying cryptographic algorithm, compared with roughly 2^n in the classical case,^[10] meaning that symmetric key lengths are effectively halved: AES-256 would have the same security against an attack using Grover's algorithm that AES-128 has against classical brute-force

search (see Key size). Quantum cryptography could potentially fulfill some of the functions of public key cryptography.

Besides factorization and discrete logarithms, quantum algorithms offering a more than polynomial speedup over the best known classical algorithm have been found for several problems,^[11] including the simulation of quantum physical processes from chemistry and solid state physics, the approximation of Jones polynomials, and solving Pell's equation. No mathematical proof has been found that shows that an equally fast classical algorithm cannot be discovered, although this is considered unlikely. For some problems, quantum computers offer a polynomial speedup. The most well-known example of this is *quantum database search*, which can be solved by Grover's algorithm using quadratically fewer queries to the database than are required by classical algorithms. In this case the advantage is provable. Several other examples of provable quantum speedups for query problems have subsequently been discovered, such as for finding collisions in two-to-one functions and evaluating NAND trees.

Consider a problem that has these four properties:

1. The only way to solve it is to guess answers repeatedly and check them,
2. There are n possible answers to check,
3. Every possible answer takes the same amount of time to check, and
4. There are no clues about which answers might be better: generating possibilities randomly is just as good as checking them in some special order.

An example of this is a password cracker that attempts to guess the password for an encrypted file (assuming that the password has a maximum possible length).

For problems with all four properties, the time for a quantum computer to solve this will be proportional to the square root of n . That can be a very large speedup, reducing some problems from years to seconds. It can be used to attack symmetric ciphers such as Triple DES and AES by attempting to guess the secret key.

Grover's algorithm can also be used to obtain a quadratic speed-up [over a brute-force search] for a class of problems known as NP-complete.

Since chemistry and nanotechnology rely on understanding quantum systems, and such systems are impossible to simulate in an efficient manner classically, many believe quantum simulation will be one of the most important applications of quantum computing.^[12]

There are a number of practical difficulties in building a quantum computer, and thus far quantum computers have only solved trivial problems. David DiVincenzo, of IBM, listed the following requirements for a practical quantum computer:^[13]

- scalable physically to increase the number of qubits;
- qubits can be initialized to arbitrary values;
- quantum gates faster than decoherence time;
- universal gate set;
- qubits can be read easily.

Quantum decoherence

One of the greatest challenges is controlling or removing quantum decoherence. This usually means isolating the system from its environment as the slightest interaction with the external world would cause the system to decohere. This effect is irreversible, as it is non-unitary, and is usually something that should be avoided, if not highly controlled. Decoherence times for candidate systems, in particular the transverse relaxation time T_2 (for NMR and MRI technology, also called the *dephasing time*), typically range between nanoseconds and seconds at low temperature.^[4]

These issues are more difficult for optical approaches as the timescales are orders of magnitude lower and an often cited approach to overcoming them is optical pulse shaping. Error rates are typically proportional to the ratio of

operating time to decoherence time, hence any operation must be completed much more quickly than the decoherence time.

If the error rate is small enough, it is thought to be possible to use quantum error correction, which corrects errors due to decoherence, thereby allowing the total calculation time to be longer than the decoherence time. An often cited figure for required error rate in each gate is 10^{-4} . This implies that each gate must be able to perform its task in one 10,000th of the decoherence time of the system.

Meeting this scalability condition is possible for a wide range of systems. However, the use of error correction brings with it the cost of a greatly increased number of required qubits. The number required to factor integers using Shor's algorithm is still polynomial, and thought to be between L and L^2 , where L is the number of bits in the number to be factored; error correction algorithms would inflate this figure by an additional factor of L . For a 1000-bit number, this implies a need for about 10^4 qubits without error correction.^[14] With error correction, the figure would rise to about 10^7 qubits. Note that computation time is about L^2 or about 10^7 steps and on 1 MHz, about 10 seconds.

A very different approach to the stability-decoherence problem is to create a topological quantum computer with anyons, quasi-particles used as threads and relying on braid theory to form stable logic gates.^{[15] [16]}

Developments

There are a number of quantum computing candidates, among those:

- Superconductor-based quantum computers (including SQUID-based quantum computers)^[17]
- Trapped ion quantum computer
- Optical lattices
- Topological quantum computer^[18]
- Quantum dot on surface (e.g. the Loss-DiVincenzo quantum computer)
- Nuclear magnetic resonance on molecules in solution (liquid NMR)
- Solid state NMR Kane quantum computers
- Electrons on helium quantum computers
- Cavity quantum electrodynamics (CQED)
- Molecular magnet
- Fullerene-based ESR quantum computer
- Optic-based quantum computers (Quantum optics)
- Diamond-based quantum computer^{[19] [20] [21]}
- Bose–Einstein condensate-based quantum computer^[22]
- Transistor-based quantum computer - string quantum computers with entrainment of positive holes using an electrostatic trap
- Spin-based quantum computer
- Adiabatic quantum computation^[23]
- Rare-earth-metal-ion-doped inorganic crystal based quantum computers^{[24] [25]}

The large number of candidates shows explicitly that the topic, in spite of rapid progress, is still in its infancy. But at the same time there is also a vast amount of flexibility.

In 2005, researchers at the University of Michigan built a semiconductor chip which functioned as an ion trap. Such devices, produced by standard lithography techniques, may point the way to scalable quantum computing tools.^[26] An improved version was made in 2006.

In 2009, researchers at Yale University created the first rudimentary solid-state quantum processor. The two-qubit superconducting chip was able to run elementary algorithms. Each of the two artificial atoms (or qubits) were made up of a billion aluminum atoms but they acted like a single one that could occupy two different energy states.^{[27] [28]}

Another team, working at the University of Bristol, also created a silicon-based quantum computing chip, based on quantum optics. The team was able to run Shor's algorithm on the chip.^[29]

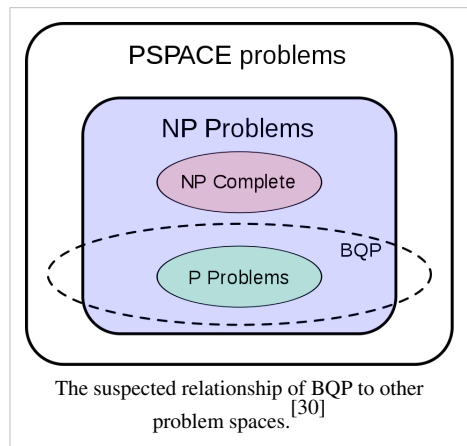
Relation to computational complexity theory

The class of problems that can be efficiently solved by quantum computers is called **BQP**, for "bounded error, quantum, polynomial time". Quantum computers only run **probabilistic** algorithms, so **BQP** on quantum computers is the counterpart of **BPP** on classical computers. It is defined as the set of problems solvable with a polynomial-time algorithm, whose probability of error is bounded away from one half.^[30] A quantum computer is said to "solve" a problem if, for every instance, its answer will be right with high probability. If that solution runs in polynomial time, then that problem is in **BQP**.

BQP is contained in the complexity class $\#P$ (or more precisely in the associated class of decision problems $P^{\#P}$),^[31] which is a subclass of PSPACE.

BQP is suspected to be disjoint from NP-complete and a strict superset of **P**, but that is not known. Both integer factorization and discrete log are in **BQP**. Both of these problems are **NP** problems suspected to be outside **BPP**, and hence outside **P**. Both are suspected to not be NP-complete. There is a common misconception that quantum computers can solve NP-complete problems in polynomial time. That is not known to be true, and is generally suspected to be false.^[31]

Although quantum computers may be faster than classical computers, those described above can't solve any problems that classical computers can't solve, given enough time and memory (however, those amounts might be practically infeasible). A Turing machine can simulate these quantum computers, so such a quantum computer could never solve an undecidable problem like the halting problem. The existence of "standard" quantum computers does not disprove the Church–Turing thesis.^[30] It has been speculated that theories of quantum gravity, such as M-theory or loop quantum gravity, may allow even faster computers to be built. Currently, it's an open problem to even *define* computation in such theories due to the problem of time, i.e. there's no obvious way to describe what it means for an observer to submit input to a computer and later receive output.^[32]



See also

- Timeline of quantum computing
- Quantum bus
- Post-quantum cryptography
- Chemical computer
- Optical computer
- DNA computer
- Molecular computer
- List of emerging technologies

References

- [1] "Quantum Computing with Molecules (<http://www.media.mit.edu/physics/publications/papers/98.06.sciam/0698gershenfeld.html>)" article in Scientific American by Neil Gershenfeld and Isaac L. Chuang - a generally accessible overview of quantum computing and so on.
- [2] Quantum Information Science and Technology Roadmap (http://qist.lanl.gov/qcomp_map.shtml) for a sense of where the research is heading.
- [3] Waldner, Jean-Baptiste (2007). *Nanocomputers and Swarm Intelligence*. London: ISTE. p. 157. ISBN 2746215160.
- [4] David P. DiVincenzo (1995). "Quantum Computation". *Science* **270** (5234): 255–261. doi:10.1126/science.270.5234.255.
- [5] Integer Factoring (http://modular.fas.harvard.edu/edu/Fall2001/124/misc/arjen_lenstra_factoring.pdf) By ARJEN K. LENSTRA - Designs, Codes and Cryptography, 19, 101–128 (2000) Kluwer Academic Publishers
- [6] Daniel J. Bernstein, Introduction to Post-Quantum Cryptography (http://pqcrypto.org/www.springer.com/cda/content/document/cda_downloadaddocument/9783540887010-c1.pdf). Introduction to Daniel J. Bernstein, Johannes Buchmann, Erik Dahmen (editors). Post-quantum cryptography. Springer, Berlin, 2009. ISBN 978-3-540-88701-0
- [7] See also pqcrypto.org (<http://pqcrypto.org/>), a bibliography maintained by Daniel J. Bernstein and Tanja Lange on cryptography not known to be broken by quantum computing.
- [8] Robert J. McEliece. "A public-key cryptosystem based on algebraic coding theory (http://ipnpr.jpl.nasa.gov/progress_report2/42-44/44N.PDF)." Jet Propulsion Laboratory DSN Progress Report 42–44, 114–116.
- [9] Kobayashi, H.; Gall, F.L. (2006), "Dihedral Hidden Subgroup Problem: A Survey", *Information and Media Technologies (J-STAGE)* **1** (1): 178–185
- [10] Bennett C.H., Bernstein E., Brassard G., Vazirani U., *The strengths and weaknesses of quantum computation* (<http://www.cs.berkeley.edu/~vazirani/pubs/bbbv.ps>). SIAM Journal on Computing 26(5): 1510–1523 (1997).
- [11] Quantum Algorithm Zoo (<http://www.its.caltech.edu/~sjordan/zoo.html>) - Stephen Jordan's Homepage
- [12] The Father of Quantum Computing (<http://www.wired.com/science/discoveries/news/2007/02/72734>) By Quinn Norton 02.15.2007, Wired.com
- [13] David P. DiVincenzo, IBM (2000-04-13). "The Physical Implementation of Quantum Computation" (<http://arxiv.org/abs/quant-ph/0002077>). . Retrieved 2006-11-17.
- [14] M. I. Dyakonov, Université Montpellier (2006-10-14). "Is Fault-Tolerant Quantum Computation Really Possible?" (<http://arxiv.org/abs/quant-ph/0610117>). . Retrieved 2007-02-16.
- [15] Freedman, Michael; Alexei Kitaev, Michael Larsen, Zhenghan Wang (2002-10-20). "Topological Quantum Computation". *Bulletin of the American Mathematical Society* **40** (1): 31–38. doi:10.1090/S0273-0979-02-00964-3.
- [16] Monroe, Don, *Anyons: The breakthrough quantum computing needs?* (<http://www.newscientist.com/channel/fundamentals/mg20026761.700-anyons-the-breakthrough-quantum-computing-needs.html>), New Scientist, 1 October 2008
- [17] Clarke, John; Wilhelm, Frank (June 19, 2008). "Superconducting quantum bits" (<http://www.nature.com/nature/journal/v453/n7198/full/nature07128.html>). *Nature* **453** (7198): 1031–1042. doi:10.1038/nature07128. ISSN 0028-0836. PMID 18563154. .
- [18] Nayak, Chetan; Simon, Steven; Stern, Ady (2008). "Nonabelian Anyons and Quantum Computation" (<http://arxiv.org/abs/0707.1889>). *Rev Mod Phys* **80**: 1083. .
- [19] Nizovtsev, A. P.; Kilin, S. Ya.; Jelezko, F.; Gaebel, T.; Popa, I.; Gruber, A.; Wrachtrup, J. (October 19, 2004). "A quantum computer based on NV centers in diamond: Optically detected nutations of single electron and nuclear spins" (<http://www.springerlink.com/content/5p6554lg35716085/>). *Optics and Spectroscopy* **99** (2): 248–260. doi:10.1134/1.2034610. .
- [20] Wolfgang Gruener, TG Daily (2007-06-01). "Research indicates diamonds could be key to quantum storage" (<http://www.tgdaily.com/content/view/32306/118/>). . Retrieved 2007-06-04.
- [21] Neumann, P.; Mizuochi, N.; Rempp, F.; Hemmer, P.; Watanabe, H.; Yamasaki, S.; Jacques, V.; Gaebel, T. *et al.* (June 6, 2008). "Multipartite Entanglement Among Single Spins in Diamond" (<http://www.sciencemag.org/cgi/content/abstract/320/5881/1326>). *Science* **320** (5881): 1326–1329. doi:10.1126/science.1157233. PMID 18535240. .
- [22] Rene Millman, IT PRO (2007-08-03). "Trapped atoms could advance quantum computing" (<http://www.itpro.co.uk/news/121086/trapped-atoms-could-advance-quantum-computing.html>). . Retrieved 2007-07-26.
- [23] William M Kaminsky, MIT (Date Unknown). "Scalable Superconducting Architecture for Adiabatic Quantum Computation" (<http://arxiv.org/pdf/quant-ph/0403090>). . Retrieved 2007-02-19.
- [24] Ohlsson, N.; Mohan, R. K.; Kröll, S. (January 1, 2002). "Quantum computer hardware based on rare-earth-ion-doped inorganic crystals" (<http://www.sciencedirect.com/science/article/B6TVF-44J3RM9-J/2/307aab59d157ddd2ebb8281f76f89138>). *Opt. Commun.* **201** (1–3): 71–77. doi:10.1016/S0030-4018(01)01666-2. .
- [25] Longdell, J. J.; Sellars, M. J.; Manson, N. B. (September 23, 2004). "Demonstration of conditional quantum phase shift between ions in a solid" (<http://prola.aps.org/abstract/PRL/v93/i13/e130503>). *Phys. Rev. Lett.* **93** (13): 130503. doi:10.1103/PhysRevLett.93.130503. PMID 15524694. .
- [26] Ann Arbor (2005-12-12). "U-M develops scalable and mass-producible quantum computer chip" (<http://www.umich.edu/news/index.html?Releases/2005/Dec05/r121205b>). . Retrieved 2006-11-17.
- [27] Dicarlo, L; Chow, JM; Gambetta, JM; Bishop, LS; Johnson, BR; Schuster, DI; Majer, J; Blais, A *et al.* (2009-06-28). "Demonstration of two-qubit algorithms with a superconducting quantum processor" (<http://www.nature.com/nature/journal/vaop/ncurrent/pdf/nature08121.pdf>). *Nature* **460** (7252): 240–4. doi:10.1038/nature08121. ISSN 0028-0836. PMID 19561592. . Retrieved 2009-07-02.

- [28] "Scientists Create First Electronic Quantum Processor" (<http://opa.yale.edu/news/article.aspx?id=6764>). 2009-07-02. . Retrieved 2009-07-02.
- [29] New Scientist (2009-09-04). "Code-breaking quantum algorithm runs on a silicon chip" (<http://www.newscientist.com/article/dn17736-codebreaking-quantum-algorithm-run-on-a-silicon-chip.html>). . Retrieved 2009-10-14.
- [30] Michael Nielsen and Isaac Chuang (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press. ISBN 0-521-63503-9. OCLC 174527496.
- [31] Bernstein and Vazirani, Quantum complexity theory, SIAM Journal on Computing, 26(5):1411-1473, 1997. (<http://www.cs.berkeley.edu/~vazirani/bv.ps>)
- [32] Scott Aaronson, *NP-complete Problems and Physical Reality* (<http://arxiv.org/abs/quant-ph/0502072>), ACM SIGACT News, Vol. 36, No. 1. (March 2005), pp. 30-52, section 7 "Quantum Gravity": "[...] to anyone who wants a test or benchmark for a favorite quantum gravity theory,[author's footnote: That is, one without all the bother of making numerical predictions and comparing them to observation] let me humbly propose the following: *can you define Quantum Gravity Polynomial-Time?* [...] until we can say what it means for a 'user' to specify an 'input' and 'later' receive an 'output'—*there is no such thing as computation, not even theoretically.*" (emphasis in original)

General references

- Derek Abbott, Charles R. Doering, Carlton M. Caves, Daniel M. Lidar, Howard E. Brandt, Alexander R. Hamilton, David K. Ferry, Julio Gea-Banacloche, Sergey M. Bezrukov, and Laszlo B. Kish (2003). "Dreams versus Reality: Plenary Debate Session on Quantum Computing". *Quantum Information Processing* **2** (6): 449–472. doi:10.1023/B:QINP.0000042203.24782.9a. arXiv:quant-ph/0310130. (Alternative Location (free) at Michigan university's repository Deep Blue (<http://hdl.handle.net/2027.42/45526>))
- David P. DiVincenzo (2000). "The Physical Implementation of Quantum Computation". *Experimental Proposals for Quantum Computation*. arXiv:quant-ph/0002077
- David P. DiVincenzo (1995). "Quantum Computation". *Science* **270** (5234): 255–261. doi:10.1126/science.270.5234.255. Table 1 lists switching and dephasing times for various systems.
- Richard Feynman (1982). "Simulating physics with computers". *International Journal of Theoretical Physics* **21**: 467. doi:10.1007/BF02650179.
- Gregg Jaeger (2006). *Quantum Information: An Overview*. Berlin: Springer. ISBN 0-387-35725-4. OCLC 255569451.
- Michael Nielsen and Isaac Chuang (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press. ISBN 0-521-63503-9. OCLC 174527496.
- Stephanie Frank Singer (2005). *Linearity, Symmetry, and Prediction in the Hydrogen Atom*. New York: Springer. ISBN 0-387-24637-1. OCLC 253709076.
- Giuliano Benenti (2004). *Principles of Quantum Computation and Information Volume 1*. New Jersey: World Scientific. ISBN 9-812-38830-3. OCLC 179950736.
- David P. DiVincenzo (2000). "The Physical Implementation of Quantum Computation". *Experimental Proposals for Quantum Computation*. arXiv:quant-ph/0002077.
- Sam Lomonaco Four Lectures on Quantum Computing given at Oxford University in July 2006 (<http://www.csee.umbc.edu/~lomonaco/Lectures.html#OxfordLectures>)
- C. Adami, N.J. Cerf. (1998). "Quantum computation with linear optics". arXiv:quant-ph/9806048v1.
- Joachim Stolze,; Dieter Suter, (2004). *Quantum Computing*. Wiley-VCH. ISBN 3527404384.
- Ian Mitchell, (1998). "Computing Power into the 21st Century: Moore's Law and Beyond" (<http://citeseer.ist.psu.edu/mitchell98computing.html>).
- Rolf Landauer, (1961). "Irreversibility and heat generation in the computing process" (<http://www.research.ibm.com/journal/rd/053/ibmrd0503C.pdf>).
- Gordon E. Moore (1965). *Cramming more components onto integrated circuits*.
- R.W. Keyes, (1988). *Miniaturization of electronics and its limits*.
- M. A. Nielsen,; E. Knill, ; R. Laflamme,., "Complete Quantum Teleportation By Nuclear Magnetic Resonance" (<http://citeseer.ist.psu.edu/595490.html>).

- Lieven M.K. Vandersypen,; Constantino S. Yannoni, ; Isaac L. Chuang, (2000). *Liquid state NMR Quantum Computing*.
- Imai Hiroshi,; Hayashi Masahito, (2006). *Quantum Computation and Information*. Berlin: Springer. ISBN 3540331328.
- Andre Berthiaume, (1997). "Quantum Computation" (<http://citeseer.ist.psu.edu/article/berthiaume97quantum.html>).
- Daniel R. Simon, (1994). "On the Power of Quantum Computation" (<http://citeseer.ist.psu.edu/simon94power.html>). Institute of Electrical and Electronic Engineers Computer Society Press.
- "Seminar Post Quantum Cryptology" (http://www.crypto.rub.de/its_seminar_ss08.html). Chair for communication security at the Ruhr-University Bochum.
- Laura Sanders, (2009). "First programmable quantum computer created" (http://www.sciencenews.org/view/generic/id/49951/title/First_programmable_quantum_computer_created).

External links

- Stanford Encyclopedia of Philosophy: " Quantum Computing (<http://plato.stanford.edu/entries/qt-quantcomp/>)" by Amit Hagar.
 - Quantiki (<http://www.quantiki.org/>) - Wiki and portal with free-content related to quantum information science.
 - jQuantum: Java quantum circuit simulator (<http://jquantum.sourceforge.net/jQuantumApplet.html>)
 - QCAD: Quantum circuit emulator (<http://www.phys.cs.is.nagoya-u.ac.jp/~watanabe/qcad/index.html>)
 - C++ Quantum Library (<https://gna.org/projects/quantumlibrary>)
 - Haskell Library for Quantum computations (<http://hackage.haskell.org/cgi-bin/hackage-scripts/package/quantum-arrow>)
 - Video Lectures by David Deutsch (http://www.quiprocone.org/Protected/DD_lectures.htm)
 - Lectures at the Institut Henri Poincaré (slides and videos) (<http://www.quantware.ups-tlse.fr/IHP2006/>)
-

Quantum chemistry

Quantum chemistry is a branch of theoretical chemistry, which applies quantum mechanics and quantum field theory to address problems in chemistry. The description of the electronic behavior of atoms and molecules as pertaining to their reactivity is one of the applications of quantum chemistry. Quantum chemistry lies on the border between chemistry and physics, and significant contributions have been made by scientists from both fields. It has a strong and active overlap with the field of atomic physics and molecular physics, as well as physical chemistry.

Quantum chemistry mathematically describes the fundamental behavior of matter at the molecular scale.^[1] It is, in principle, possible to describe all chemical systems using this theory. In practice, only the simplest chemical systems may realistically be investigated in purely quantum mechanical terms, and approximations must be made for most practical purposes (e.g., Hartree-Fock, post Hartree-Fock or Density functional theory, see computational chemistry for more details). Hence a detailed understanding of quantum mechanics is not necessary for most chemistry, as the important implications of the theory (principally the orbital approximation) can be understood and applied in simpler terms.

In quantum mechanics the Hamiltonian, or the physical state, of a particle can be expressed as the sum of two operators, one corresponding to kinetic energy and the other to potential energy. The Hamiltonian in the Schrödinger wave equation used in quantum chemistry does not contain terms for the spin of the electron.

Solutions of the Schrödinger equation for the hydrogen atom gives the form of the wave function for atomic orbitals, and the relative energy of the various orbitals. The orbital approximation can be used to understand the other atoms e.g. helium, lithium and carbon.

History

The history of quantum chemistry essentially began with the 1838 discovery of cathode rays by Michael Faraday, the 1859 statement of the black body radiation problem by Gustav Kirchhoff, the 1877 suggestion by Ludwig Boltzmann that the energy states of a physical system could be discrete, and the 1900 quantum hypothesis by Max Planck that any energy radiating atomic system can theoretically be divided into a number of discrete energy elements ϵ such that each of these energy elements is proportional to the frequency ν with which they each individually radiate energy, as defined by the following formula:

$$\epsilon = h\nu$$

where h is a numerical value called Planck's Constant. Then, in 1905, to explain the photoelectric effect (1839), i.e., that shining light on certain materials can function to eject electrons from the material, Albert Einstein postulated, based on Planck's quantum hypothesis, that light itself consists of individual quantum particles, which later came to be called photons (1926). In the years to follow, this theoretical basis slowly began to be applied to chemical structure, reactivity, and bonding.

Electronic structure

The first step in solving a quantum chemical problem is usually solving the Schrödinger equation (or Dirac equation in relativistic quantum chemistry) with the electronic molecular Hamiltonian. This is called determining the **electronic structure** of the molecule. It can be said that the electronic structure of a molecule or crystal implies essentially its chemical properties. An exact solution for the Schrödinger equation can only be obtained for the hydrogen atom. Since all other atomic, or molecular systems, involve the motions of three or more "particles", their Schrödinger equations cannot be solved exactly and so approximate solutions must be sought.

Wave model

The foundation of quantum mechanics and quantum chemistry is the **wave model**, in which the atom is a small, dense, positively charged nucleus surrounded by electrons. Unlike the earlier Bohr model of the atom, however, the wave model describes electrons as "clouds" moving in orbitals, and their positions are represented by probability distributions rather than discrete points. The strength of this model lies in its predictive power. Specifically, it predicts the pattern of chemically similar elements found in the periodic table. The wave model is so named because electrons exhibit properties (such as interference) traditionally associated with waves. See wave-particle duality.

Valence bond

Although the mathematical basis of quantum chemistry had been laid by Schrödinger in 1926, it is generally accepted that the first true calculation in quantum chemistry was that of the German physicists Walter Heitler and Fritz London on the hydrogen (H_2) molecule in 1927. Heitler and London's method was extended by the American theoretical physicist John C. Slater and the American theoretical chemist Linus Pauling to become the **Valence-Bond (VB)** [or **Heitler-London-Slater-Pauling (HLSP)**] method. In this method, attention is primarily devoted to the pairwise interactions between atoms, and this method therefore correlates closely with classical chemists' drawings of bonds.

Molecular orbital

An alternative approach was developed in 1929 by Friedrich Hund and Robert S. Mulliken, in which electrons are described by mathematical functions delocalized over an entire molecule. The **Hund-Mulliken** approach or **molecular orbital (MO) method** is less intuitive to chemists, but has turned out capable of predicting spectroscopic properties better than the VB method. This approach is the conceptional basis of the **Hartree-Fock method** and further post Hartree-Fock methods.

Density functional theory

The **Thomas-Fermi model** was developed independently by Thomas and Fermi in 1927. This was the first attempt to describe many-electron systems on the basis of electronic density instead of wave functions, although it was not very successful in the treatment of entire molecules. The method did provide the basis for what is now known as **density functional theory**. Though this method is less developed than post Hartree-Fock methods, its significantly lower computational requirements (scaling typically no worse than n^3 with respect to n basis functions) allow it to tackle larger polyatomic molecules and even macromolecules. This computational affordability and often comparable accuracy to MP2 and CCSD (post-Hartree-Fock methods) has made it one of the most popular methods in computational chemistry at present.

Chemical dynamics

A further step can consist of solving the Schrödinger equation with the total molecular Hamiltonian in order to study the motion of molecules. Direct solution of the Schrödinger equation is called *quantum molecular dynamics*, within the semiclassical approximation *semiclassical molecular dynamics*, and within the classical mechanics framework *molecular dynamics (MD)*. Statistical approaches, using for example Monte Carlo methods, are also possible.

Adiabatic chemical dynamics

In **adiabatic dynamics**, interatomic interactions are represented by single scalar potentials called potential energy surfaces. This is the Born-Oppenheimer approximation introduced by Born and Oppenheimer in 1927. Pioneering applications of this in chemistry were performed by Rice and Ramsperger in 1927 and Kassel in 1928, and generalized into the RRKM theory in 1952 by Marcus who took the transition state theory developed by Eyring in

1935 into account. These methods enable simple estimates of unimolecular reaction rates from a few characteristics of the potential surface.

Non-adiabatic chemical dynamics

Non-adiabatic dynamics consists of taking the interaction between several coupled potential energy surface (corresponding to different electronic quantum states of the molecule). The coupling terms are called **vibronic couplings**. The pioneering work in this field was done by Stueckelberg, Landau, and Zener in the 1930s, in their work on what is now known as the Landau-Zener transition. Their formula allows the transition probability between two diabatic potential curves in the neighborhood of an avoided crossing to be calculated.

Quantum chemistry and quantum field theory

The application of quantum field theory (QFT) to chemical systems and theories has become increasingly common in the modern physical sciences. One of the first and most fundamentally explicit appearances of this is seen in the theory of the photomagnetron. In this system, plasmas, which are ubiquitous in both physics and chemistry, are studied in order to determine the basic quantization of the underlying bosonic field. However, quantum field theory is of interest in many fields of chemistry, including: nuclear chemistry, astrochemistry, sonochemistry, and quantum hydrodynamics. Field theoretic methods have also been critical in developing the ab initio Effective Hamiltonian theory of semi-empirical pi-electron methods.

See also

- Atomic physics
- Computational chemistry
- Condensed matter physics
- International Academy of Quantum Molecular Science
- Molecular modelling
- Physical chemistry
- Quantum chemistry computer programs
- Quantum electrochemistry
- QMC@Home
- Theoretical physics

Further reading

- Atkins, P.W. Friedman, R. (2005). *Molecular Quantum Mechanics*, Oxford University Press, 4th edition. ISBN 978-0199274987
- Atkins, P.W. *Physical Chemistry* (Oxford University Press) ISBN 0-19-879285-9
- Atkins, P.W. Friedman, R. (2008). *Quanta, Matter and Change: A Molecular Approach to Physical Change*, W. H. Freeman. ISBN 978-0716761174
- Bernard Pullman and Alberte Pullman. 1963. *Quantum Biochemistry*, New York and London: Academic Press.
- Eric R. Scerri, The Periodic Table: Its Story and Its Significance, Oxford University Press, 2006. Considers the extent to which chemistry and especially the periodic system has been reduced to quantum mechanics. ISBN 0-19-530573-6.
- McWeeny, R. *Coulson's Valence* (Oxford Science Publications) ISBN 0-19-855144-4
- Karplus M., Porter R.N. (1971). *Atoms and Molecules. An introduction for students of physical chemistry*, Benjamin-Cummings Publishing Company, ISBN 978-0805352184

- Landau, L.D. and Lifshitz, E.M. *Quantum Mechanics:Non-relativistic Theory* (Course of Theoretical Physics vol.3) (Pergamon Press)
- Levine, I. (2008). *Physical Chemistry*, McGraw-Hill Science, 6th edition. ISBN 978-0072538625 (Hardcover) or ISBN 978-0071276368 (Paperback)
- Pauling, L. (1954). *General Chemistry*. Dover Publications. ISBN 0-486-65622-5.
- Pauling, L., and Wilson, E. B. (1935/1963). *Introduction to Quantum Mechanics with Applications to Chemistry* (Dover Publications) ISBN 0-486-64871-0
- Simon, Z. (1976). *Quantum Biochemistry and Specific Interactions.*, Taylor & Francis; ISBN 978-0856260872 and ISBN 0-85-6260878 .

External links

- The Sherrill Group - Notes ^[2]
- ChemViz Curriculum Support Resources ^[3]
- Early ideas in the history of quantum chemistry ^[4]
- The Particle Adventure ^[5]

Nobel lectures by quantum chemists

- Walter Kohn's Nobel lecture ^[6]
- Rudolph Marcus' Nobel lecture ^[7]
- Robert Mulliken's Nobel lecture ^[8]
- Linus Pauling's Nobel lecture ^[9]
- John Pople's Nobel lecture ^[10]

References

- [1] "Quantum Chemistry" (http://cmm.cit.nih.gov/modeling/guide_documents/quantum_mechanics_document.html). *The NIH Guide to Molecular Modeling*. National Institutes of Health. . Retrieved 2007-09-08.
 - [2] <http://vergil.chemistry.gatech.edu/notes/index.html>
 - [3] <http://www.shodor.org/chemviz/>
 - [4] <http://www.quantum-chemistry-history.com/>
 - [5] <http://particleadventure.org/>
 - [6] <http://nobelprize.org/chemistry/laureates/1998/kohn-lecture.html>
 - [7] <http://nobelprize.org/chemistry/laureates/1992/marcus-lecture.html>
 - [8] <http://nobelprize.org/chemistry/laureates/1966/mulliken-lecture.html>
 - [9] <http://nobelprize.org/chemistry/laureates/1954/pauling-lecture.html>
 - [10] <http://nobelprize.org/chemistry/laureates/1998/pople-lecture.html>
-

Density functional theory

Electronic structure methods
Tight binding
Nearly-free electron model
Hartree–Fock
Modern valence bond
Generalized valence bond
Møller–Plesset perturbation theory
Configuration interaction
Coupled cluster
Multi-configurational self-consistent field
Density functional theory
Quantum chemistry composite methods
Quantum Monte Carlo
k·p perturbation theory
Muffin-tin approximation
LCAO method

Density functional theory (DFT) is a quantum mechanical theory used in physics and chemistry to investigate the electronic structure (principally the ground state) of many-body systems, in particular atoms, molecules, and the condensed phases. With this theory, the properties of a many-electron system can be determined by using functionals, i.e. functions of another function, which in this case is the spatially dependent electron density. Hence the name density functional theory comes from the use of functionals of the electron density. DFT is among the most popular and versatile methods available in condensed-matter physics, computational physics, and computational chemistry.

DFT has been very popular for calculations in solid state physics since the 1970s. In many cases the results of DFT calculations for solid-state systems agreed quite satisfactorily with experimental data. Also, the computational costs were relatively low when compared to traditional ways which were based on the complicated many-electron wavefunction, such as Hartree-Fock theory and its descendants. However, DFT was not considered accurate enough for calculations in quantum chemistry until the 1990s, when the approximations used in the theory were greatly refined to better model the exchange and correlation interactions. DFT is now a leading method for electronic structure calculations in chemistry and solid-state physics.

Despite the improvements in DFT, there are still difficulties in using density functional theory to properly describe intermolecular interactions, especially van der Waals forces (dispersion); charge transfer excitations; transition states, global potential energy surfaces and some other strongly correlated systems; and in calculations of the band gap in semiconductors. Its poor treatment of dispersion renders DFT unsuitable (at least when used alone) for the treatment of systems which are dominated by dispersion (e.g., interacting noble gas atoms) or where dispersion competes significantly with other effects (e.g. in biomolecules). The development of new DFT methods designed to overcome this problem, by alterations to the functional or by the inclusion of additive terms, is a current research topic.

Overview of method

Although density functional theory has its conceptual roots in the Thomas-Fermi model, DFT was put on a firm theoretical footing by the two **Hohenberg-Kohn theorems** (H-K).^[1] The original H-K theorems held only for non-degenerate ground states in the absence of a magnetic field, although they have since been generalized to encompass these.^{[2] [3]}

The first H-K theorem demonstrates that the ground state properties of a many-electron system are uniquely determined by an electron density that depends on only 3 spatial coordinates. It lays the groundwork for reducing the many-body problem of N electrons with $3N$ spatial coordinates to 3 spatial coordinates, through the use of functionals of the electron density. This theorem can be extended to the time-dependent domain to develop time-dependent density functional theory (TDDFT), which can be used to describe excited states.

The second H-K theorem defines an energy functional for the system and proves that the correct ground state electron density minimizes this energy functional.

Within the framework of Kohn-Sham DFT, the intractable many-body problem of interacting electrons in a static external potential is reduced to a tractable problem of non-interacting electrons moving in an effective potential. The effective potential includes the external potential and the effects of the Coulomb interactions between the electrons, e.g., the exchange and correlation interactions. Modeling the latter two interactions becomes the difficulty within KS DFT. The simplest approximation is the local-density approximation (LDA), which is based upon exact exchange energy for a uniform electron gas, which can be obtained from the Thomas-Fermi model, and from fits to the correlation energy for a uniform electron gas. Non-interacting systems are relatively easy to solve as the wavefunction can be represented as a Slater determinant of orbitals. Further, the kinetic energy functional of such a system is known exactly. The exchange-correlation part of the total-energy functional remains unknown and must be approximated.

Another approach, less popular than Kohn-Sham DFT (KS-DFT) but arguably more closely related to the spirit of the original H-K theorems, is orbital-free density functional theory (OFDFT), in which approximate functionals are also used for the kinetic energy of the non-interacting system.

Derivation and formalism

As usual in many-body electronic structure calculations, the nuclei of the treated molecules or clusters are seen as fixed (the Born-Oppenheimer approximation), generating a static external potential V in which the electrons are moving. A stationary electronic state is then described by a wavefunction $\Psi(\vec{r}_1, \dots, \vec{r}_N)$ satisfying the many-electron Schrödinger equation

$$\hat{H}\Psi = \left[\hat{T} + \hat{V} + \hat{U} \right] \Psi = \left[\sum_i^N -\frac{\hbar^2}{2m} \nabla_i^2 + \sum_i^N V(\vec{r}_i) + \sum_{i<j}^N U(\vec{r}_i, \vec{r}_j) \right] \Psi = E\Psi$$

where $\hat{H}$ is the electronic molecular Hamiltonian, N is the number of electrons, $\hat{T}$ is the N -electron kinetic energy, $\hat{V}$ is the N -electron potential energy from the external field, and $\hat{U}$ is the electron-electron interaction energy for the N -electron system. The operators $\hat{T}$ and $\hat{U}$ are so-called universal operators as they are the same for any system, while $\hat{V}$ is system dependent, i.e. non-universal. The difference between having separable single-particle problems and the much more complicated many-particle problem arises from the interaction term $\hat{U}$. There are many sophisticated methods for solving the many-body Schrödinger equation based on the expansion of the wavefunction in Slater determinants. While the simplest one is the Hartree-Fock method, more sophisticated approaches are usually categorized as post-Hartree-Fock methods. However, the problem with these methods is the huge computational effort, which makes it virtually impossible to apply them efficiently to larger, more complex systems.

Here DFT provides an appealing alternative, being much more versatile as it provides a way to systematically map the many-body problem, with $\hat{U}$, onto a single-body problem without $\hat{U}$. In DFT the key variable is the particle density $n(\vec{r})$, which for a normalized Ψ is given by

$$n(\vec{r}) = N \int d^3r_2 \int d^3r_3 \cdots \int d^3r_N \Psi^*(\vec{r}, \vec{r}_2, \dots, \vec{r}_N) \Psi(\vec{r}, \vec{r}_2, \dots, \vec{r}_N)$$

This relation can be reversed, i.e. for a given ground-state density $n_0(\vec{r})$ it is possible, in principle, to calculate the corresponding ground-state wavefunction $\Psi_0(\vec{r}_1, \dots, \vec{r}_N)$. In other words, Ψ_0 is a unique functional of n_0 ,^[1]

$$\Psi_0 = \Psi[n_0]$$

and consequently the ground-state expectation value of an observable $\hat{O}$ is also a functional of n_0

$$O[n_0] = \langle \Psi[n_0] | \hat{O} | \Psi[n_0] \rangle$$

In particular, the ground-state energy is a functional of n_0

$$E_0 = E[n_0] = \langle \Psi[n_0] | \hat{T} + \hat{V} + \hat{U} | \Psi[n_0] \rangle$$

where the contribution of the external potential $\langle \Psi[n_0] | \hat{V} | \Psi[n_0] \rangle$ can be written explicitly in terms of the ground-state density n_0

$$V[n_0] = \int V(\vec{r}) n_0(\vec{r}) d^3r$$

More generally, the contribution of the external potential $\langle \Psi | \hat{V} | \Psi \rangle$ can be written explicitly in terms of the density n ,

$$V[n] = \int V(\vec{r}) n(\vec{r}) d^3r$$

The functionals $T[n]$ and $U[n]$ are called universal functionals, while $V[n]$ is called a non-universal functional, as it depends on the system under study. Having specified a system, i.e., having specified $\hat{V}$, one then has to minimize the functional

$$E[n] = T[n] + U[n] + \int V(\vec{r}) n(\vec{r}) d^3r$$

with respect to $n(\vec{r})$, assuming one has got reliable expressions for $T[n]$ and $U[n]$. A successful minimization of the energy functional will yield the ground-state density n_0 and thus all other ground-state observables.

The variational problems of minimizing the energy functional $E[n]$ can be solved by applying the Lagrangian method of undetermined multipliers.^[4] First, one considers an energy functional that doesn't explicitly have an electron-electron interaction energy term,

$$E_s[n] = \langle \Psi_s[n] | \hat{T}_s + \hat{V}_s | \Psi_s[n] \rangle$$

where $\hat{T}_s$ denotes the non-interacting kinetic energy and $\hat{V}_s$ is an external effective potential in which the particles are moving. Obviously, $n_s(\vec{r}) \stackrel{\text{def}}{=} n(\vec{r})$ if $\hat{V}_s$ is chosen to be

$$\hat{V}_s = \hat{V} + \hat{U} + (\hat{T} - \hat{T}_s)$$

Thus, one can solve the so-called Kohn-Sham equations of this auxiliary non-interacting system,

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + V_s(\vec{r}) \right] \phi_i(\vec{r}) = \epsilon_i \phi_i(\vec{r})$$

which yields the orbitals ϕ_i that reproduce the density $n(\vec{r})$ of the original many-body system

$$n(\vec{r}) \stackrel{\text{def}}{=} n_s(\vec{r}) = \sum_i^N |\phi_i(\vec{r})|^2$$

The effective single-particle potential can be written in more detail as

$$V_s(\vec{r}) = V(\vec{r}) + \int \frac{e^2 n_s(\vec{r}')}{|\vec{r} - \vec{r}'|} d^3 r' + V_{XC}[n_s(\vec{r})]$$

where the second term denotes the so-called Hartree term describing the electron-electron Coulomb repulsion, while the last term V_{XC} is called the exchange-correlation potential. Here, V_{XC} includes all the many-particle interactions. Since the Hartree term and V_{XC} depend on $n(\vec{r})$, which depends on the ϕ_i , which in turn depend on V_s , the problem of solving the Kohn-Sham equation has to be done in a self-consistent (i.e., iterative) way. Usually one starts with an initial guess for $n(\vec{r})$, then calculates the corresponding V_s and solves the Kohn-Sham equations for the ϕ_i . From these one calculates a new density and starts again. This procedure is then repeated until convergence is reached.

Approximations (Exchange-correlation functionals)

The major problem with DFT is that the exact functionals for exchange and correlation are not known except for the free electron gas. However, approximations exist which permit the calculation of certain physical quantities quite accurately. In physics the most widely used approximation is the local-density approximation (LDA), where the functional depends only on the density at the coordinate where the functional is evaluated:

$$E_{XC}[n] = \int \epsilon_{XC}(n) n(\vec{r}) d^3 r.$$

The local spin-density approximation (LSDA) is a straightforward generalization of the LDA to include electron spin:

$$E_{XC}[n_\uparrow, n_\downarrow] = \int \epsilon_{XC}(n_\uparrow, n_\downarrow) n(\vec{r}) d^3 r.$$

Highly accurate formulae for the exchange-correlation energy density $\epsilon_{XC}(n_\uparrow, n_\downarrow)$ have been constructed from quantum Monte Carlo simulations of a free electron model.^[5]

Generalized gradient approximations (GGA) are still local but also take into account the gradient of the density at the same coordinate:

$$E_{XC}[n_\uparrow, n_\downarrow] = \int \epsilon_{XC}(n_\uparrow, n_\downarrow, \vec{\nabla} n_\uparrow, \vec{\nabla} n_\downarrow) n(\vec{r}) d^3 r.$$

Using the latter (GGA) very good results for molecular geometries and ground-state energies have been achieved.

Potentially more accurate than the GGA functionals are meta-GGA functions. These functionals include a further term in the expansion, depending on the density, the gradient of the density and the Laplacian (second derivative) of the density.

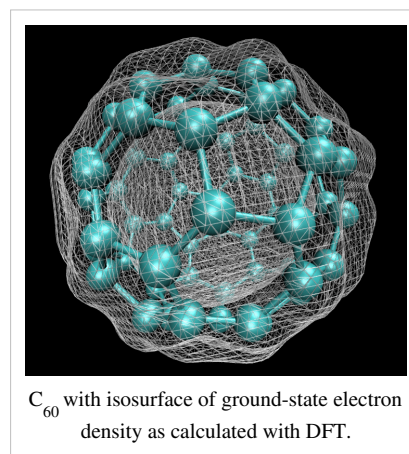
Difficulties in expressing the exchange part of the energy can be relieved by including a component of the exact exchange energy calculated from Hartree-Fock theory. Functionals of this type are known as hybrid functionals.

Generalizations to include magnetic fields

The DFT formalism described above breaks down, to various degrees, in the presence of a vector potential, i.e. a magnetic field. In such a situation, the one-to-one mapping between the ground-state electron density and wavefunction is lost. Generalizations to include the effects of magnetic fields have led to two different theories: current density functional theory (CDFT) and magnetic field functional theory (BDFT). In both these theories, the functional used for the exchange and correlation must be generalized to include more than just the electron density. In current density functional theory, developed by Vignale and Rasolt,^[3] the functionals become dependent on both the electron density and the paramagnetic current density. In magnetic field density functional theory, developed by Salsbury, Grayce and Harris, the functionals depend on the electron density and the magnetic field, and the functional form can depend on the form of the magnetic field. In both of these theories it has been difficult to develop functionals beyond their equivalent to LDA, which are also readily implementable computationally.

Applications

In practice, Kohn-Sham theory can be applied in several distinct ways depending on what is being investigated. In solid state calculations, the local density approximations are still commonly used along with plane wave basis sets, as an electron gas approach is more appropriate for electrons delocalised through an infinite solid. In molecular calculations, however, more sophisticated functionals are needed, and a huge variety of exchange-correlation functionals have been developed for chemical applications. Some of these are inconsistent with the uniform electron gas approximation, however, they must reduce to LDA in the electron gas limit. Among physicists, probably the most widely used functional is the revised Perdew-Burke-Ernzerhof exchange model (a direct generalized-gradient parametrization of the free electron gas with no free parameters); however, this is not sufficiently calorimetrically accurate for gas-phase molecular calculations. In the chemistry community, one popular functional is known as BLYP (from the name Becke for the exchange part and Lee, Yang and Parr for the correlation part). Even more widely used is B3LYP which is a hybrid functional in which the exchange energy, in this case from Becke's exchange functional, is combined with the exact energy from Hartree-Fock theory. Along with the component exchange and correlation functionals, three parameters define the hybrid functional, specifying how much of the exact exchange is mixed in. The adjustable parameters in hybrid functionals are generally fitted to a 'training set' of molecules. Unfortunately, although the results obtained with these functionals are usually sufficiently accurate for most applications, there is no systematic way of improving them (in contrast to some of the traditional wavefunction-based methods like configuration interaction or coupled cluster theory). Hence in the current DFT approach it is not possible to estimate the error of the calculations without comparing them to other methods or experiments.



For molecular applications, in particular for hybrid functionals, Kohn-Sham DFT methods are usually implemented just like Hartree-Fock itself.

Thomas–Fermi model

The predecessor to density functional theory was the **Thomas–Fermi model**, developed by Thomas and Fermi in 1927. They used a statistical model to approximate the distribution of electrons in an atom. The mathematical basis postulated that electrons are distributed uniformly in phase space with two electrons in every h^3 of volume.^[6] For each element of coordinate space volume d^3r we can fill out a sphere of momentum space up to the Fermi momentum p_f ^[7]

$$(4/3)\pi p_f^3(\vec{r})$$

Equating the number of electrons in coordinate space to that in phase space gives:

$$n(\vec{r}) = \frac{8\pi}{3h^3} p_f^3(\vec{r})$$

Solving for p_f and substituting into the classical kinetic energy formula then leads directly to a kinetic energy represented as a functional of the electron density:

$$t_{TF}[n] = \frac{p^2}{2m_e} \propto \frac{(n^{1/3})^2}{2m_e} \propto n^{2/3}(\vec{r})$$

$$T_{TF}[n] = C_F \int n(\vec{r}) n^{2/3}(\vec{r}) d^3r = C_F \int n^{5/3}(\vec{r}) d^3r$$

$$\text{where } C_F = \frac{3h^2}{10m_e} \left(\frac{3}{8\pi} \right)^{2/3}$$

As such, they were able to calculate the energy of an atom using this kinetic energy functional combined with the classical expressions for the nuclear-electron and electron-electron interactions (which can both also be represented in terms of the electron density).

Although this was an important first step, the Thomas–Fermi equation's accuracy is limited because the resulting kinetic energy functional is only approximate, and because the method does not attempt to represent the exchange energy of an atom as a conclusion of the Pauli principle. An exchange energy functional was added by Dirac in 1928.

However, the Thomas–Fermi–Dirac theory remained rather inaccurate for most applications. The largest source of error was in the representation of the kinetic energy, followed by the errors in the exchange energy, and due to the complete neglect of electron correlation.

Teller (1962) showed that Thomas–Fermi theory cannot describe molecular bonding. This can be overcome by improving the kinetic energy functional.

The kinetic energy functional can be improved by adding the Weizsäcker (1935) correction.^{[8] [9]}

$$T_W[n] = \frac{1}{8} \frac{\hbar^2}{m} \int \frac{|\nabla n(\vec{r})|^2}{n(\vec{r})} d\vec{r}$$

Hohenberg-Kohn Theorem

1. For N -interacting electrons, $E[n]$ is only functional of the electron density.

2. $E[n_{GS}] = E_{GS}$

E_{GS} is the real ground state energy, and n_{GS} is the real ground state electron density.

Software supporting DFT

DFT is supported by many Quantum chemistry and solid state physics codes, often along with other methods.

See also

- Basis set (chemistry)
- Gas in a box
- Helium atom
- Kohn–Sham equations
- Local density approximation
- Molecule
- Molecular design software
- Molecular modelling
- Quantum chemistry
- List of quantum chemistry and solid state physics software
- List of software for molecular mechanics modeling
- Thomas–Fermi model
- Time-dependent density functional theory

Books on DFT

- R. Dreizler, E. Gross, Density Functional Theory (Plenum Press, New York, 1995).
- C. Fiolhais, F. Nogueira, M. Marques (eds.), A Primer in Density Functional Theory (Springer-Verlag, 2003). [10]
- Kohanoff, J., Electronic Structure Calculations for Solids and Molecules: Theory and Computational Methods (Cambridge University Press, 2006).
- W. Koch, M. C. Holthausen, A Chemist's Guide to Density Functional Theory (Wiley-VCH, Weinheim, ed. 2, 2002).
- R. G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules (Oxford University Press, New York, 1989), ISBN 0-19-504279-4, ISBN 0-19-509276-7 (pbk.).
- N.H. March, Electron Density Theory of Atoms and Molecules (Academic Press, 1992), ISBN 0-12-470525-1.
- Richard M. Martin, Electronic Structure: Basic Theory and Practical Methods, Cambridge University Press, 2004
- D. Sholl, J. A. Steckel, Density Functional Theory: A Practical Introduction, Wiley-Interscience, 2009

Key papers

- L.H. Thomas, The calculation of atomic fields, Proc. Camb. Phil. Soc, 23 542-548
- P. Hohenberg and W. Kohn, Phys. Rev. 136 (1964) B864 [11]
- W. Kohn and L. J. Sham, Phys. Rev. 140 (1965) A1133 [12]
- A. D. Becke, J. Chem. Phys. 98 (1993) 5648 [13]
- C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B 37 (1988) 785 [14]
- P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. 98 (1994) 11623 [15]
- K. Burke, J. Werschnik, and E. K. U. Gross, *Time-dependent density functional theory: Past, present, and future*. J. Chem. Phys. 123, 062206 [16] (2005). OAI: arXiv.org:cond-mat/0410362 [17].

External links

- Walter Kohn, Nobel Laureate [18] Freeview video interview with Walter on his work developing density functional theory by the Vega Science Trust.
- Klaus Capelle, A bird's-eye view of density-functional theory [19]
- Walter Kohn, Nobel Lecture [20]
- Density functional theory on arxiv.org [21]
- FreeScience Library -> Density Functional Theory [22]
- The ABC of DFT [23]
- Density Functional Theory -- an introduction [24]
- Electron Density Functional Theory - Lecture Notes [25]

References

- [1] Hohenberg, Pierre; Walter Kohn (1964). "Inhomogeneous electron gas". *Physical Review* **136** (3B): B864–B871. doi:10.1103/PhysRev.136.B864.
- [2] Levy, Mel (1979). "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v -representability problem". *Proceedings of the National Academy of Sciences* (United States National Academy of Sciences) **76** (12): 6062–6065. doi:10.1073/pnas.76.12.6062.
- [3] Vignale, G.; Mark Rasolt (1987). "Density-functional theory in strong magnetic fields". *Physical Review Letters* (American Physical Society) **59** (20): 2360–2363. doi:10.1103/PhysRevLett.59.2360.
- [4] Kohn, W.; Sham, L. J. (1965). "Self-consistent equations including exchange and correlation effects". *Phys. Rev.* **140** (4A): A1133–A1138. doi:10.1103/PhysRev.140.A1133.
- [5] John P. Perdew, Adrienn Ruzsinszky, Jianmin Tao, Viktor N. Staroverov, Gustavo Scuseria and Gábor I. Csonka (2005). "Prescriptions for the design and selection of density functional approximations: More constraint satisfaction with fewer fits". *J. Chem. Phys.* **123**: 062201. doi:10.1063/1.1904565.

- [6] Parr and Yang 1989, p.47
- [7] March 1992, p.24
- [8] Weizsäcker, C. F. v. (1935). "Zur Theorie der Kernmassen". *Zeitschrift für Physik* **96** (7-8): 431–58. doi:10.1007/BF01337700.
- [9] Parr and Yang 1989, p.127
- [10] <http://www.springerlink.com/content/jyahrga7a129/>
- [11] http://prola.aps.org/abstract/PR/v136/i3B/pB864_1
- [12] http://prola.aps.org/abstract/PR/v140/i4A/pA1133_1
- [13] <http://dx.doi.org/10.1063/1.464913>
- [14] http://prola.aps.org/abstract/PRB/v37/i2/p785_1
- [15] <http://dx.doi.org/10.1021/j100096a001>
- [16] <http://dx.doi.org/10.1063/1.1904586>
- [17] <http://arxiv.org/abs/cond-mat/0410362>
- [18] <http://www.vega.org.uk/video/programme/23>
- [19] <http://arxiv.org/abs/cond-mat/0211443>
- [20] <http://nobelprize.org/chemistry/laureates/1998/kohn-lecture.pdf>
- [21] <http://xstructure.inr.ac.ru/x-bin/theme3.py?level=1&index1=447765>
- [22] <http://freescience.info/books.php?id=30>
- [23] <http://chem.ps.uci.edu/~kieron/dft/book/gamma/g1.pdf>
- [24] <http://arxiv.org/abs/physics/9806013>
- [25] http://www.fh.huji.ac.il/~roib/LectureNotes/DFT/DFT_Course_Roi_Baer.pdf

Birefringence

Birefringence, or **double refraction**, is the decomposition of a ray of light into two rays (the **ordinary ray** and the **extraordinary ray**) when it passes through certain types of material, such as calcite crystals or boron nitride, depending on the polarization of the light. This effect can occur only if the structure of the material is anisotropic (directionally dependent). If the material has a single axis of anisotropy



A calcite crystal laid upon a paper with some letters showing the double refraction

or optical axis, (i.e. it is uniaxial) birefringence can be formalized by assigning two different refractive indices to the material for different polarizations. The birefringence magnitude is then defined by

$$\Delta n = n_e - n_o$$

where n_e and n_o are the refractive indices for polarizations parallel (**extraordinary**) and perpendicular (**ordinary**) to the axis of anisotropy respectively.^[1]

The reason for birefringence is the fact that in anisotropic media the electric field vector $\vec{E}$ and the dielectric displacement $\vec{D}$ can be nonparallel (namely for the extraordinary polarisation), although being linearly related.

Birefringence can also arise in magnetic, not dielectric, materials, but substantial variations in magnetic permeability of materials are rare at optical frequencies. Liquid crystal materials as used in Liquid Crystal Displays (LCDs) are also birefringent.^[2]

Creation

While birefringence is often found naturally (especially in crystals), there are several ways to create it in optically isotropic materials.

- Birefringence results when isotropic materials are deformed such that the isotropy is lost in one direction (i.e., stretched or bent). Example ^[3]
- Applying an electric field can induce molecules to line up or behave asymmetrically, introducing anisotropy and resulting in birefringence. (*see* Pockels effect)
- Applying a magnetic field can cause a material to be **circularly birefringent**, with different indices of refraction for oppositely-handed circular polarizations
- Self alignment of highly polar molecules such as lipids and some surfactants will generate highly birefringent thin films (see also Liquid crystal)

Examples of uniaxial birefringent materials

Uniaxial materials, at 590 nm^[4]

Material	n_o	n_e	Δn
beryl $\text{Be}_3\text{Al}_2(\text{SiO}_3)_6$	1.602	1.557	-0.045
calcite CaCO_3	1.658	1.486	-0.172
calomel Hg_2Cl_2	1.973	2.656	+0.683
ice H_2O	1.309	1.313	+0.004
lithium niobate LiNbO_3	2.272	2.187	-0.085
magnesium fluoride MgF_2	1.380	1.385	+0.006
quartz SiO_2	1.544	1.553	+0.009
ruby Al_2O_3	1.770	1.762	-0.008
rutile TiO_2	2.616	2.903	+0.287
peridot $(\text{Mg}, \text{Fe})_2\text{SiO}_4$	1.690	1.654	-0.036
sapphire Al_2O_3	1.768	1.760	-0.008
sodium nitrate NaNO_3	1.587	1.336	-0.251
tourmaline (complex silicate)	1.669	1.638	-0.031
zircon, high ZrSiO_4	1.960	2.015	+0.055
zircon, low ZrSiO_4	1.920	1.967	+0.047

Many plastics are birefringent, because their molecules are 'frozen' in a stretched conformation when the plastic is moulded or extruded.^[5] For example, cellophane is a cheap birefringent material, and Polaroid sheets are commonly used to examine for orientation in birefringent plastics like polystyrene and polycarbonate. Birefringent materials are used in many devices which manipulate the polarization of light, such as wave plates, polarizing prisms, and Lyot filters.

There are many birefringent crystals: birefringence was first described in calcite crystals by the Danish scientist Rasmus Bartholin in 1669.

Birefringence can be observed in amyloid plaque deposits such as are found in the brains of Alzheimer's patients. Modified proteins such as immunoglobulin light chains abnormally accumulate between cells, forming fibrils. Multiple folds of these fibers line up and take on a beta-pleated sheet conformation. Congo red dye intercalates

between the folds and, when observed under polarized light, causes birefringence.

Cotton (*Gossypium hirsutum*) fiber is birefringent because of high levels of cellulosic material in the fiber's secondary cell wall.

Slight imperfections in optical fiber can cause birefringence, which can cause distortion in fiber-optic communication; see polarization mode dispersion. The imperfections can be geometrically based, or a result of photoelastic effects from loading on the optical fiber.

Silicon carbide, also known as Moissanite, is strongly birefringent.

The refractive indices of several (uniaxial) birefringent materials are listed below (at wavelength ~ 590 nm)^[4]

Biaxial birefringence

Biaxial materials, at 590 nm^[4]

Material	n_{α}	n_{β}	n_{γ}
borax	1.447	1.469	1.472
epsom salt $\text{MgSO}_4 \cdot 7(\text{H}_2\text{O})$	1.433	1.455	1.461
mica, biotite	1.595	1.640	1.640
mica, muscovite	1.563	1.596	1.601
olivine $(\text{Mg, Fe})_2\text{SiO}_4$	1.640	1.660	1.680
perovskite CaTiO_3	2.300	2.340	2.380
topaz	1.618	1.620	1.627
ulexite	1.490	1.510	1.520

Biaxial birefringence, also known as **trirefringence**, describes an anisotropic material that has more than one axis of anisotropy. For such a material, the refractive index tensor $\mathbf{n}$, will in general have three distinct eigenvalues that can be labeled n_{α} , n_{β} and n_{γ} .

Measurement

Birefringence and related optical effects (such as optical rotation and linear or circular dichroism) can be measured by measuring the changes in the polarization of light passing through the material. These measurements are known as polarimetry.

Birefringence of lipid bilayers can be measured using dual polarisation interferometry. This provides a measure of the degree of order within these fluid layers and how this order is disrupted when the layer interacts with other biomolecules.

A common feature of optical microscopes is a pair of crossed polarizing filters. Between the crossed polarizers, a birefringent sample will appear bright against a dark (isotropic) background.

For a fixed composition such as calcium carbonate, a crystal such as calcite or its polymorphs, the index of refraction depends on the direction of light through the crystal structure. The refraction also depends on composition, and can be calculated using the Gladstone-Dale relation.

Applications

Birefringence is widely used in optical devices, such as liquid crystal displays, light modulators, color filters, wave plates, optical axis gratings, etc. It also plays an important role in second harmonic generation and many other nonlinear processes. It is also utilized in medical diagnostics: needle aspiration of fluid from a gouty joint will reveal negatively birefringent urate crystals. In ophthalmology, scanning laser polarimetry utilises the birefringence of the retinal nerve fibre layer to indirectly quantify its thickness, which is of use in the assessment and monitoring of glaucoma. Birefringence characteristics in sperm heads allow for the selection of spermatozoa for intracytoplasmic sperm injection.^[6]

Birefringent filters are also used as spatial low-pass filters in electronic cameras, where the thickness of the crystal is controlled to spread the image in one direction, thus increasing the spot-size. This is essential to the proper working of all television and electronic film cameras, to avoid spatial aliasing, the folding back of frequencies higher than can be sustained by the pixel matrix of the camera.

Elastic birefringence

Another form of birefringence is observed in anisotropic elastic materials. In these materials, shear waves split according to similar principles as the light waves discussed above. The study of birefringent shear waves in the earth is a part of seismology. Birefringence is also used in optical mineralogy to determine the chemical composition, and history of minerals and rocks.

Electromagnetic waves in an anisotropic material

Effective refractive indices in uniaxial materials

Propagation direction	Ordinary ray		Extraordinary ray	
	Polarization	n_{eff}	Polarization	n_{eff}
z	xy -plane	n_o	n/a	n/a
xy -plane	xy -plane	n_o	z	n_e
xz -plane	y	n_o	xz -plane	$n_e < n < n_o$
other	analogous to xz -plane			

The behavior of a light ray that propagates through an anisotropic material is dependent on its polarization. For a given propagation direction, there are generally two perpendicular polarizations for which the medium behaves as if it had a single effective refractive index. In a uniaxial material, rays with these polarizations are called the extraordinary and the ordinary ray (e and o rays), corresponding to the extraordinary and ordinary refractive indices. In a biaxial material, there are three refractive indices α , β , and γ , yet only two rays, which are called the fast and the slow ray. The slow ray is the ray that has the highest effective refractive index.

For a uniaxial material with the z axis defined to be the optical axis, the effective refractive indices are as in the table on the right. For rays propagating in the xz plane, the effective refractive index of the e polarization varies continuously between n_o and n_e , depending on the angle with the z axis. The effective refractive index can be constructed from the Index ellipsoid.

Mathematical description

More generally, birefringence can be defined by considering a dielectric permittivity and a refractive index that are tensors. Consider a plane wave propagating in an anisotropic medium, with a relative permittivity tensor ϵ , where the refractive index $\mathbf{n}$, is defined by $\mathbf{n} \cdot \mathbf{n} = \epsilon$. If the wave has an electric vector of the form:

$$\mathbf{E} = \mathbf{E}_0 \exp i(\mathbf{k} \cdot \mathbf{r} - \omega t) \quad (2)$$

where $\mathbf{r}$ is the position vector and t is time, then the wave vector $\mathbf{k}$ and the angular frequency ω must satisfy Maxwell's equations in the medium, leading to the equations:

$$-\nabla \times \nabla \times \mathbf{E} = \frac{1}{c^2} (\epsilon \cdot \frac{\partial^2 \mathbf{E}}{\partial t^2}) \quad (3a)$$

$$\nabla \cdot (\epsilon \cdot \mathbf{E}) = 0 \quad (3b)$$

where c is the speed of light in a vacuum. Substituting eqn. 2 in eqns. 3a-b leads to the conditions:

$$|\mathbf{k}|^2 \mathbf{E}_0 - (\mathbf{k} \cdot \mathbf{E}_0) \mathbf{k} = \frac{\omega^2}{c^2} (\epsilon \cdot \mathbf{E}_0) \quad (4a)$$

$$\mathbf{k} \cdot (\epsilon \cdot \mathbf{E}_0) = 0 \quad (4b)$$

For the matrix product $(\epsilon \cdot \mathbf{E})$ often a separate name is used, the *dielectric displacement vector* $\mathbf{D}$. So essentially birefringence concerns the general theory of linear relationships between these two vectors in anisotropic media. To find the allowed values of $\mathbf{k}$, $\mathbf{E}_0$ can be eliminated from eq 4a. One way to do this is to write eqn 4a in Cartesian coordinates, where the x , y and z axes are chosen in the directions of the eigenvectors of ϵ , so that

$$\epsilon = \begin{bmatrix} n_x^2 & 0 & 0 \\ 0 & n_y^2 & 0 \\ 0 & 0 & n_z^2 \end{bmatrix} \quad (4c)$$

Hence eqn 4a becomes

$$(-k_y^2 - k_z^2 + \frac{\omega^2 n_x^2}{c^2}) E_x + k_x k_y E_y + k_x k_z E_z = 0 \quad (5a)$$

$$k_x k_y E_x + (-k_x^2 - k_z^2 + \frac{\omega^2 n_y^2}{c^2}) E_y + k_y k_z E_z = 0 \quad (5b)$$

$$k_x k_z E_x + k_y k_z E_y + (-k_x^2 - k_y^2 + \frac{\omega^2 n_z^2}{c^2}) E_z = 0 \quad (5c)$$

where E_x, E_y, E_z, k_x, k_y and k_z are the components of $\mathbf{E}_0$ and $\mathbf{k}$. This is a set of linear equations in E_x, E_y, E_z , and they have a non-trivial solution if their determinant is zero:

$$\det \begin{bmatrix} (-k_y^2 - k_z^2 + \frac{\omega^2 n_x^2}{c^2}) & k_x k_y & k_x k_z \\ k_x k_y & (-k_x^2 - k_z^2 + \frac{\omega^2 n_y^2}{c^2}) & k_y k_z \\ k_x k_z & k_y k_z & (-k_x^2 - k_y^2 + \frac{\omega^2 n_z^2}{c^2}) \end{bmatrix} = 0 \quad (6)$$

Multiplying out eqn (6), and rearranging the terms, we obtain

$$\frac{\omega^4}{c^4} - \frac{\omega^2}{c^2} \left(\frac{k_x^2 + k_y^2}{n_z^2} + \frac{k_x^2 + k_z^2}{n_y^2} + \frac{k_y^2 + k_z^2}{n_x^2} \right) + \left(\frac{k_x^2}{n_y^2 n_z^2} + \frac{k_y^2}{n_x^2 n_z^2} + \frac{k_z^2}{n_x^2 n_y^2} \right) (k_x^2 + k_y^2 + k_z^2) = 0 \quad (7)$$

In the case of a uniaxial material, where $n_x = n_y = n_o$ and $n_z = n_e$ say, eqn 7 can be factorised into

$$\left(\frac{k_x^2}{n_o^2} + \frac{k_y^2}{n_o^2} + \frac{k_z^2}{n_o^2} - \frac{\omega^2}{c^2} \right) \left(\frac{k_x^2}{n_e^2} + \frac{k_y^2}{n_e^2} + \frac{k_z^2}{n_o^2} - \frac{\omega^2}{c^2} \right) = 0 \quad (8)$$

Each of the factors in eqn 8 defines a surface in the space of vectors $\mathbf{k}$ — the **surface of wave normals**. The first factor defines a sphere and the second defines an ellipsoid. Therefore, for each direction of the wave normal, two wavevectors $\mathbf{k}$ are allowed. Values of $\mathbf{k}$ on the sphere correspond to the **ordinary rays** while values on the ellipsoid correspond to the **extraordinary rays**.

For a biaxial material, eqn (7) cannot be factorized in the same way, and describes a more complicated pair of wave-normal surfaces.^[7]

Birefringence is often measured for rays propagating along one of the optical axes (or measured in a two-dimensional material). In this case, $\mathbf{n}$ has two eigenvalues which can be labeled n_1 and n_2 . $\mathbf{n}$ can be diagonalized by:

$$\mathbf{n} = \mathbf{R}(\chi) \cdot \begin{bmatrix} n_1 & 0 \\ 0 & n_2 \end{bmatrix} \cdot \mathbf{R}(\chi)^T \quad (9)$$

where $\mathbf{R}(\chi)$ is the rotation matrix through an angle χ . Rather than specifying the complete tensor $\mathbf{n}$, we may now simply specify the *magnitude* of the birefringence Δn , and *extinction angle* χ , where $\Delta n = n_1 - n_2$.

See also

- Cotton-Mouton effect
- Crystal optics
- John Kerr
- Periodic poling
- Dichroism

External links

- <http://www.olympusmicro.com/primer/lightandcolor/birefringence.html>
- [8] Video of stress birefringence in Polymethylmethacrylate (PMMA or Plexiglas).
- Application note on the theory of birefringence ^[9] (see no.14)

References

- [1] Eric Weisstein's World of Science on Birefringence (<http://scienceworld.wolfram.com/physics/Birefringence.html>)
- [2] The Science of Color, by Steven K. Shevell, Optical Society of America. Published 2003. ISBN 0444512519
- [3] <http://www.oberlin.edu/physics/catalog/demonstrations/optics/birefringence.html>
- [4] Elert, Glenn. "Refraction" (<http://hypertextbook.com/physics/waves/refraction/>). *The Physics Hypertextbook*. .
- [5] The Use of Birefringence for Predicting the Stiffness of Injection Moulded Polycarbonate Discs (http://www.dep.uminho.pt/home/rec_humanos/mostra_curriculum.php3?pessoa=12&&menu=5&&idcategoria=1)
- [6] Gianaroli L, Magli MC, Ferraretti AP, *et al.* (December 2008). "Birefringence characteristics in sperm heads allow for the selection of reacted spermatozoa for intracytoplasmic sperm injection". *Fertil. Steril.*. doi:10.1016/j.fertnstert.2008.10.024. PMID 19064263.
- [7] Born M, and Wolf E, *Principles of Optics*, 7th Ed. 1999 (Cambridge University Press), §15.3.3
- [8] <http://www.youtube.com/watch?v=BEC1YQbuG7U>
- [9] http://www.campoly.com/application_notes.html

Polarization spectroscopy

Polarization spectroscopy comprises a set of spectroscopic techniques based on polarization properties of light (not necessarily visible one; UV, X-ray, infrared, or in any other frequency range of the electromagnetic radiation). By analyzing the polarization properties of light, decisions can be made about the media that emitted the light (or the media the light passes/scatters through). Alternatively, a source of polarized light may be used to probe a media; in this case, the changes in the light polarization (comparing to the incidental one) allow to infer the media properties.

In general, any kind of anisotropy in the media results in some sort of light polarization. Such an anisotropy can be either inherent to the media (e.g., in the case of a crystal substance), or imposed externally (e.g., in the presence of magnetic field in plasma).

See also:

- Zeeman effect
- Faraday effect
- Stark effect
- Plasma diagnostics

Polarized IR Spectroscopy

Infrared spectroscopy (IR spectroscopy) is the subset of spectroscopy that deals with the infrared region of the electromagnetic spectrum. It covers a range of techniques, the most common being a form of absorption spectroscopy. As with all spectroscopic techniques, it can be used to identify compounds and investigate sample composition. A common laboratory instrument that uses this technique is an infrared spectrophotometer.

The infrared portion of the electromagnetic spectrum is usually divided into three regions; the near-, mid- and far-infrared, named for their relation to the visible spectrum. The far-infrared, approximately $400\text{--}10\text{ cm}^{-1}$ ($1000\text{--}30\text{ }\mu\text{m}$), lying adjacent to the microwave region, has low energy and may be used for rotational spectroscopy. The mid-infrared, approximately $4000\text{--}400\text{ cm}^{-1}$ ($30\text{--}2.5\text{ }\mu\text{m}$) may be used to study the fundamental vibrations and associated rotational-vibrational structure. The higher energy near-IR, approximately $14000\text{--}4000\text{ cm}^{-1}$ ($2.5\text{--}0.8\text{ }\mu\text{m}$) can excite overtone or harmonic vibrations. The names and classifications of these subregions are merely conventions. They are neither strict divisions nor based on exact molecular or electromagnetic properties.

Theory

Infrared spectroscopy exploits the fact that molecules absorb specific frequencies that are characteristic of their structure. These absorptions are resonant frequencies, i.e. the frequency of the absorbed radiation matches the frequency of the bond or group that vibrates. The energies are determined by the shape of the molecular potential energy surfaces, the masses of the atoms, and the associated vibronic coupling.

In particular, in the Born–Oppenheimer and harmonic approximations, i.e. when the molecular Hamiltonian corresponding to the electronic ground state can be approximated by a harmonic oscillator in the neighborhood of the equilibrium molecular geometry, the resonant frequencies are determined by the normal modes corresponding to the molecular electronic ground state potential energy surface. Nevertheless, the resonant frequencies can be in a first approach related to the strength of the bond, and the mass of the atoms at either end of it. Thus, the frequency of the vibrations can be associated with a particular bond type.

Number of vibrational modes

In order for a vibrational mode in a molecule to be "IR active," it must be associated with changes in the permanent dipole.

A molecule can vibrate in many ways, and each way is called a *vibrational mode*. Linear molecules have $3N-5$ degrees of vibrational modes whereas nonlinear molecules have $3N-6$ degrees of vibrational modes (also called vibrational degrees of freedom). As an example H_2O , a non-linear molecule, will have $3 \times 3 - 6 = 3$ degrees of vibrational freedom, or modes.

Simple diatomic molecules have only one bond and only one vibrational band. If the molecule is symmetrical, e.g. N_2 , the band is not observed in the IR spectrum, but only in the Raman spectrum. Unsymmetrical diatomic molecules, e.g. CO , absorb in the IR spectrum. More complex molecules have many bonds, and their vibrational spectra are correspondingly more complex, i.e. big molecules have many peaks in their IR spectra.

The atoms in a CH_2 group, commonly found in organic compounds can vibrate in six different ways: **symmetrical and antisymmetrical stretching, scissoring, rocking, wagging and twisting**:

Symmetrical stretching	Antisymmetrical stretching	Scissoring	Rocking	Wagging	Twisting
					

Special effects

The simplest and most important IR bands arise from the "normal modes," the simplest distortions of the molecule. In some cases, "overtone bands" are observed. These bands arise from the absorption of a photon that leads to a doubly excited vibrational state. Such bands appear at approximately twice the energy of the normal mode. Some vibrations, so-called 'combination modes,' involve more than one normal mode. The phenomenon of Fermi resonance can arise when two modes are similar in energy, Fermi resonance results in an unexpected shift in energy and intensity of the bands.

Practical IR spectroscopy

The infrared spectrum of a sample is recorded by passing a beam of infrared light through the sample. Examination of the transmitted light reveals how much energy was absorbed at each wavelength. This can be done with a monochromatic beam, which changes in wavelength over time, or by using a Fourier transform instrument to measure all wavelengths at once. From this, a transmittance or absorbance spectrum can be produced, showing at which IR wavelengths the sample absorbs. Analysis of these absorption characteristics reveals details about the molecular structure of the sample. When the frequency of the IR is the same as the vibrational frequency of a bond, absorption occurs.

This technique works almost exclusively on samples with covalent bonds. Simple spectra are obtained from samples with few IR active bonds and high levels of purity. More complex molecular structures lead to more absorption bands and more complex spectra. The technique has been used for the characterization of very complex mixtures.

Sample preparation

Gaseous samples require a sample cell with a long pathlength (typically 5–10 cm), to compensate for the diluteness.

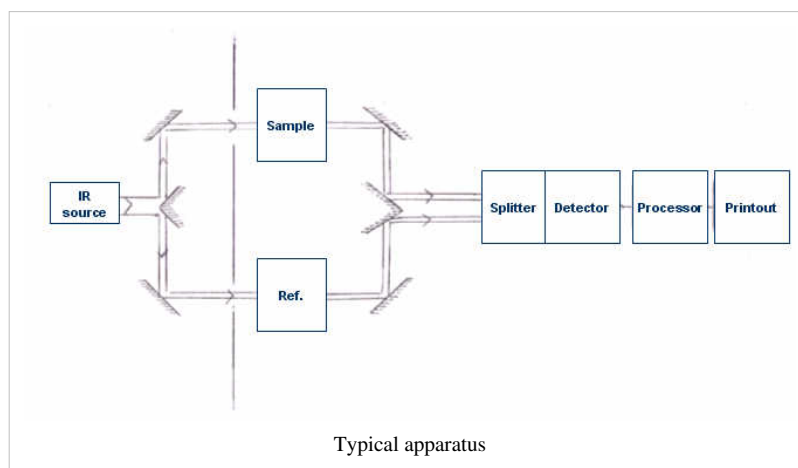
Liquid samples can be sandwiched between two plates of a salt (commonly sodium chloride, or common salt, although a number of other salts such as potassium bromide or calcium fluoride are also used).^[1] The plates are transparent to the infrared light and do not introduce any lines onto the spectra.

Solid samples can be prepared in a variety of ways. One common method is to crush the sample with an oily mulling agent (usually Nujol) in a marble or agate mortar, with a pestle. A thin film of the mull is smeared onto salt plates and measured. The second method is to grind a quantity of the sample with a specially purified salt (usually potassium bromide) finely (to remove scattering effects from large crystals). This powder mixture is then pressed in a mechanical die press to form a translucent pellet through which the beam of the spectrometer can pass.^[1] A third technique is the "cast film" technique, which is used mainly for polymeric materials. The sample is first dissolved in a suitable, non hygroscopic solvent. A drop of this solution is deposited on surface of KBr or NaCl cell. The solution is then evaporated to dryness and the film formed on the cell is analysed directly. Care is important to ensure that the film is not too thick otherwise light cannot pass through. This technique is suitable for qualitative analysis. The final method is to use microtomy to cut a thin (20–100 micrometre) film from a solid sample. This is one of the most important ways of analysing failed plastic products for example because the integrity of the solid is preserved.

It is important to note that spectra obtained from different sample preparation methods will look slightly different from each other due to differences in the samples' physical states.

Conventional apparatus

A beam of infrared light is produced and split into two separate beams. One is passed through the sample, the other passed through a reference which is often the substance the sample is dissolved in. The beams are both reflected back towards a detector, however first they pass through a splitter which quickly alternates which of the two beams enters the detector. The two signals are then compared and a printout is obtained.



A reference prevents fluctuations in the output of the source affecting the data. The reference also allows the effects of the solvent to be cancelled out (the reference is usually a pure form of the solvent the sample is in)

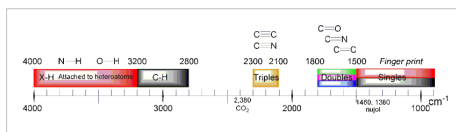
FT-IR method

Fourier transform infrared (FTIR) spectroscopy is a measurement technique for collecting infrared spectra. Instead of recording the amount of energy absorbed when the frequency of the infra-red light is varied (monochromator), the IR light is guided through an interferometer. After passing through the sample, the measured signal is the interferogram. Performing a Fourier transform on this signal data results in a spectrum identical to that from conventional (dispersive) infrared spectroscopy.

FTIR spectrometers are cheaper than conventional spectrometers because building an interferometer is easier than the fabrication of a monochromator. In addition, measurement of a single spectrum is faster for the FTIR technique because the information at all frequencies is collected simultaneously. This allows multiple samples to be collected and averaged together resulting in an improvement in sensitivity. Virtually all modern infrared spectrometers are

FTIR instruments.

Absorptions bands



Wavenumbers listed in cm^{-1} .

Uses and applications

Infrared spectroscopy is widely used in both research and industry as a simple and reliable technique for measurement, quality control and dynamic measurement. It is of especial use in forensic analysis in both criminal and civil cases, enabling identification of polymer degradation for example.

The instruments are now small, and can be transported, even for use in field trials. With increasing technology in computer filtering and manipulation of the results, samples in solution can now be measured accurately (water produces a broad absorbance across the range of interest, and thus renders the spectra unreadable without this computer treatment). Some instruments will also automatically tell you what substance is being measured from a store of thousands of reference spectra held in storage.

By measuring at a specific frequency over time, changes in the character or quantity of a particular bond can be measured. This is especially useful in measuring the degree of polymerization in polymer manufacture. Modern research instruments can take infrared measurements across the whole range of interest as frequently as 32 times a second. This can be done whilst simultaneous measurements are made using other techniques. This makes the observations of chemical reactions and processes quicker and more accurate.

Techniques have been developed to assess the quality of tea-leaves using infrared spectroscopy. This will mean that highly trained experts (also called 'noses') can be used more sparingly, at a significant cost saving.^[2]

Infrared spectroscopy has been highly successful for applications in both organic and inorganic chemistry. Infrared spectroscopy has also been successfully utilized in the field of semiconductor microelectronics^[3]: for example, infrared spectroscopy can be applied to semiconductors like silicon, gallium arsenide, gallium nitride, zinc selenide, amorphous silicon, silicon nitride, etc.

Isotope effects

The different isotopes in a particular species may give fine detail in infrared spectroscopy. For example, the O-O stretching frequency (in reciprocal centimeters) of oxyhemocyanin is experimentally determined to be 832 and 788 cm^{-1} for $\nu(^{16}\text{O}-^{16}\text{O})$ and $\nu(^{18}\text{O}-^{18}\text{O})$ respectively.

By considering the O-O as a spring, the wavenumber of absorbance, ν can be calculated:

$$\nu = \frac{1}{2\pi c} \sqrt{\frac{k}{\mu}}$$

where k is the spring constant for the bond, c is the speed of light, and μ is the reduced mass of the A-B system:

$$\mu = \frac{m_A m_B}{m_A + m_B}$$

(m_i is the mass of atom i).

The reduced masses for $^{16}\text{O}-^{16}\text{O}$ and $^{18}\text{O}-^{18}\text{O}$ can be approximated as 8 and 9 respectively. Thus

$$\frac{\nu_{16O}}{\nu_{18O}} = \sqrt{\frac{9}{8}} \approx \frac{832}{788}.$$

Where ν is the wavenumber [wavenumber = frequency/(speed of light)]

The effect of isotopes, both on the vibration and the decay dynamics, has been found to be stronger than previously thought. In some systems, such as silicon and germanium, the decay of the anti-symmetric stretch mode of interstitial oxygen involves the symmetric stretch mode with a strong isotope dependence. For example, it was shown that for a natural silicon sample, the lifetime of the anti-symmetric vibration is 11.4 ps. When the isotope of one of the silicon atoms is increased to ^{29}Si , the lifetime increases to 19 ps, similarly, when the silicon atom is changed to ^{30}Si , the lifetime becomes 27 ps.^[4]

Two-dimensional IR

Two-dimensional infrared correlation spectroscopy analysis is the application of 2D correlation analysis on infrared spectra. By extending the spectral information of a perturbed sample, spectral analysis is simplified and resolution is enhanced. The 2D synchronous and 2D asynchronous spectra represent a graphical overview of the spectral changes due to a perturbation (such as a changing concentration or changing temperature) as well as the relationship between the spectral changes at two different wavenumbers.

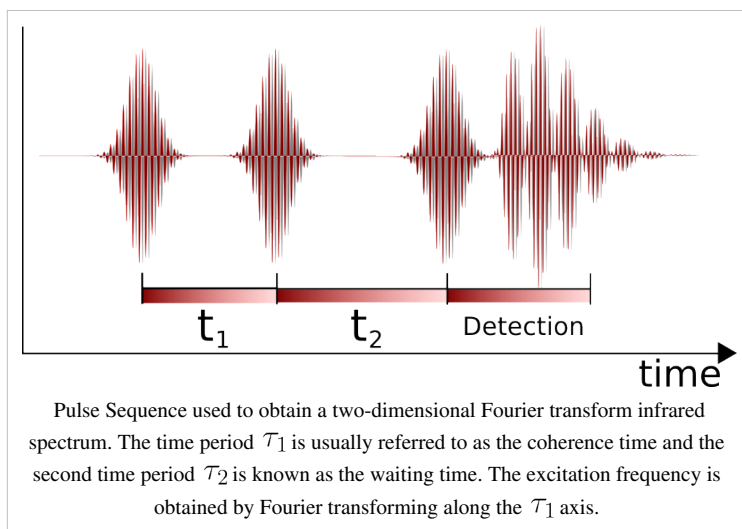
Nonlinear two-dimensional infrared spectroscopy^{[5] [6]}

is the infrared version of correlation spectroscopy.

Nonlinear two-dimensional infrared spectroscopy is a technique that has become available with the development of femtosecond infrared laser pulses. In this experiment first a set of pump pulses are applied to the sample. This is followed by a waiting time, where the system is allowed to relax. The waiting time typically lasts from zero to several picoseconds and the duration can be controlled with a resolution of tens of femtoseconds. A probe pulse is then applied

resulting in the emission of a signal from the sample. The nonlinear two-dimensional infrared spectrum is a two-dimensional correlation plot of the frequency ω_1 that was excited by the initial pump pulses and the frequency ω_3 excited by the probe pulse after the waiting time. This allows the observation of coupling between different vibrational modes; because of its extremely high time resolution it can be used to monitor molecular dynamics on a picosecond timescale. It is still a largely unexplored technique and is becoming increasingly popular for fundamental research.

Like in two-dimensional nuclear magnetic resonance (2DNMR) spectroscopy this technique spreads the spectrum in two dimensions and allow for the observation of cross peaks that contain information on the coupling between different modes. In contrast to 2DNMR nonlinear two-dimensional infrared spectroscopy also involve the excitation to overtones. These excitations result in excited state absorption peaks located below the diagonal and cross peaks. In 2DNMR two distinct techniques, COSY and NOESY, are frequently used. The cross peaks in the first are related to the scalar coupling, while in the later they are related to the spin transfer between different nuclei. In nonlinear two-dimensional infrared spectroscopy analogs have been drawn to these 2DNMR techniques. Nonlinear two-dimensional infrared spectroscopy with zero waiting time corresponds to COSY and nonlinear two-dimensional infrared spectroscopy with finite waiting time allowing vibrational population transfer corresponds to NOESY. The



COSY variant of nonlinear two-dimensional infrared spectroscopy has been used for determination of the secondary structure content proteins.^[7]

See also

- Infrared spectroscopy correlation table
- Fourier transform spectroscopy
- Near infrared spectroscopy
- Vibrational spectroscopy
- Rotational spectroscopy
- Time-resolved spectroscopy
- Spectroscopy
- Quantum vibration
- Raman spectroscopy
- Infrared microscopy
- Photothermal microspectroscopy
- Polymer degradation
- Infrared astronomy
- Far infrared astronomy
- Forensic chemistry
- Forensic engineering
- Forensic polymer engineering
- Forensic science
- Applied spectroscopy

External links

- A useful gif animation of different vibrational modes: here^[8]
- Infrared spectroscopy for organic chemists^[9]
- Organic compounds spectrum database^[10]

References

- [1] Laurence M. Harwood, Christopher J. Moody. *Experimental organic chemistry: Principles and Practice* (Illustrated edition ed.). pp. 292.
- [2] Luypaert, J.; Zhang, M.H.; Massart, D.L. (2003), "Feasibility study for the use of near infrared spectroscopy in the qualitative and quantitative analysis of green tea, *Camellia sinensis* (L.)", *Analytica Chimica Acta*, **478**(2), Elsevier, pp. 303–312
- [3] Lau, W.S. (1999). *Infrared characterization for microelectronics*. World Scientific.
- [4] Isotope Dependence of the Lifetime of the 1136-cm[sup -1] Vibration of Oxygen in Silicon K. K. Kohli, Gordon Davies, N. Q. Vinh, D. West, S. K. Estreicher, T. Gregorkiewicz, I. Izeddin, and K. M. Itoh, Phys. Rev. Lett. 96, 225503 (2006), DOI:10.1103/PhysRevLett.96.225503
- [5] P. Hamm, M. H. Lim, R. M. Hochstrasser (1998). "Structure of the amide I band of peptides measured by femtosecond nonlinear-infrared spectroscopy". *J. Phys. Chem. B* **102**: 6123. doi:10.1021/jp9813286.
- [6] S. Mukamel (2000). "Multidimensional Femtosecond Correlation Spectroscopies of Electronic and Vibrational Excitations". *Annual Review of Physics and Chemistry* **51**: 691. doi:10.1146/annurev.physchem.51.1.691.
- [7] N. Demirdöven, C. M. Cheatum, H. S. Chung, M. Khalil, J. Knoester, A. Tokmakoff (2004). "Two-dimensional infrared spectroscopy of antiparallel beta-sheet secondary structure". *Journal of the American Chemical Society* **126**: 7981. doi:10.1021/ja049811j.
- [8] <http://www.shu.ac.uk/schools/sci/chem/tutorials/molspec/irspec1.htm>
- [9] <http://www.organicworldwide.net/infrared>
- [10] http://riodb01.ibase.aist.go.jp/sdbs/cgi-bin/cre_index.cgi?lang=eng

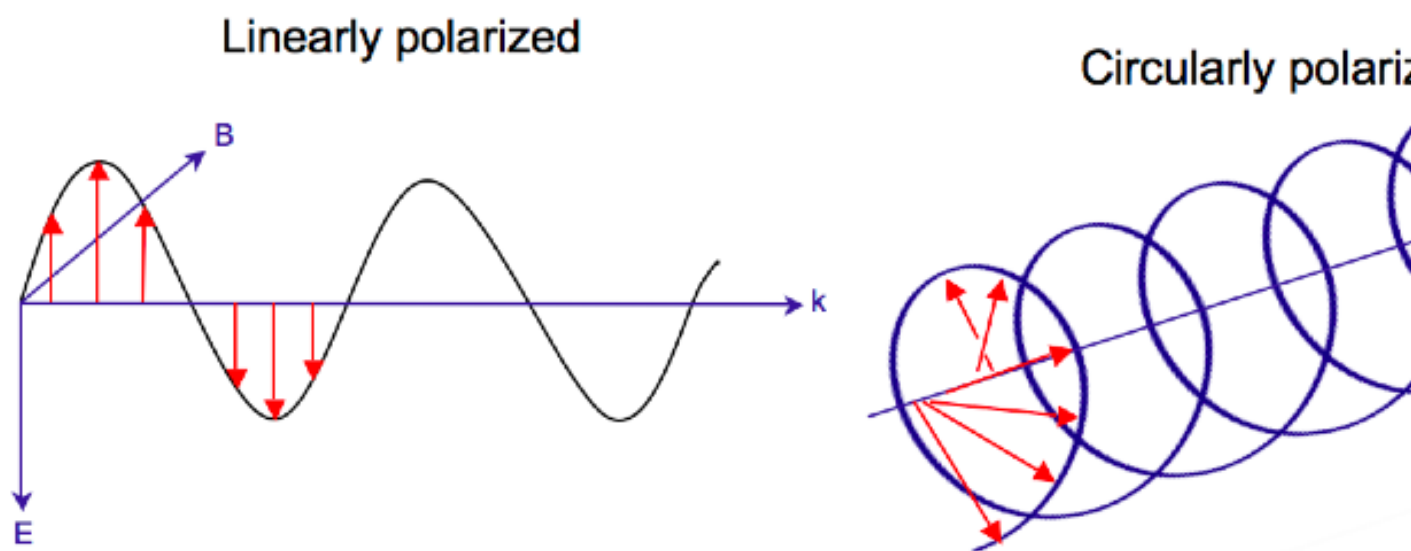
Circular dichroism

First pioneered by Jean-Baptiste Biot, Augustin Fresnel, and Aimé Cotton^[1], **circular dichroism (CD)** refers to the differential absorption of left and right circularly polarized light.^{[2] [3]} This phenomenon is exhibited in the absorption bands of optically active chiral molecules. CD spectroscopy has a wide range of applications in many different fields. Most notably, UV CD is used to investigate the secondary structure of proteins^[4]. UV/Vis CD is used to investigate charge-transfer transitions^[5]. Near-infrared CD is used to investigate geometric and electronic structure by probing metal d→d transitions^[6]. Vibrational circular dichroism, which uses light from the infrared energy region, is used for structural studies of small organic molecules, and most recently proteins and DNA^[7].

Physical Principles

Circular polarization of light

Electromagnetic radiation consists of an electric and magnetic field that oscillate perpendicular to one another and to the propagating direction^[8]. While linearly polarized light occurs when the electric field vector oscillates only in one plane and changes in magnitude, circularly polarized light occurs when the electric field vector rotates about its propagation direction and retains constant magnitude. Hence, it forms a helix in space while propagating. For left circularly polarized light (LCP) with propagation towards the observer, the electric vector rotates counterclockwise^[9]. For right circularly polarized light (RCP), the electric vector rotates clockwise.



Interaction of circularly polarized light with matter

When circularly polarized light passes through an absorbing optically active medium, the speeds between right and left polarizations differ ($c_L \neq c_R$) as well as their wavelength ($\lambda_L \neq \lambda_R$) and the extent to which they are absorbed ($\epsilon_L \neq \epsilon_R$). *Circular dichroism* is the difference $\Delta\epsilon \equiv \epsilon_L - \epsilon_R$ ^[10]. The electric field of a light beam causes a linear displacement of charge when interacting with a molecule (electric dipole), whereas the magnetic field of it causes a circulation of charge (magnetic dipole). These two motions combined cause an excitation of an electron in a helical motion, which includes translation and rotation and their associated operators. The experimentally determined relationship between the rotational strength (R) of a sample and the $\Delta\epsilon$ is given by

$$R_{exp} = \frac{3hc10^3 \ln(10)}{32\pi^3 N_A} \int \frac{\Delta\epsilon}{\nu} d\nu$$

The rotational strength has also been determined theoretically,

$$R_{theo} = \frac{1}{2mc} Im \int \Psi_g \widehat{M}_{(elec.dipole)} \Psi_e d\tau \bullet \int \Psi_g \widehat{M}_{(mag.dipole)} \Psi_e d\tau$$

We see from these two equations that in order to have non-zero $\Delta\epsilon$, the electric and magnetic dipole moment operators ($\widehat{M}_{(elec.dipole)}$ and $\widehat{M}_{(mag.dipole)}$) must transform as the same irreducible representation. C_n and D_n are the only point groups where this can occur, making only chiral molecules CD active.

Simply put, since circularly polarized light itself is "chiral", it interacts differently with chiral molecules. That is, the two types of circularly polarized light are absorbed to different extents. In a CD experiment, equal amounts of left and right circularly polarized light of a selected wavelength are alternately radiated into a (chiral) sample. One of the two polarizations is absorbed more than the other one, and this wavelength-dependent difference of absorption is measured, yielding the CD spectrum of the sample. Due to the interaction with the molecule, the electric field vector of the light traces out an elliptical path after passing through the sample.

Delta absorbance

By definition,

$$\Delta A = A_L - A_R$$

where ΔA (Delta Absorbance) is the difference between absorbance of left circularly polarized (LCP) and right circularly polarized (RCP) light (this is what is usually measured). ΔA is a function of wavelength, so for a measurement to be meaningful the wavelength it was performed at must be known.

Molar circular dichroism

It can also be expressed, by applying Beer's law, as:

$$\Delta A = (\epsilon_L - \epsilon_R)Cl$$

where

ϵ_L and ϵ_R are the molar extinction coefficients for LCP and RCP light,

C is the molar concentration

l is the path length in centimeters (cm).

Then

$$\Delta\epsilon = \epsilon_L - \epsilon_R$$

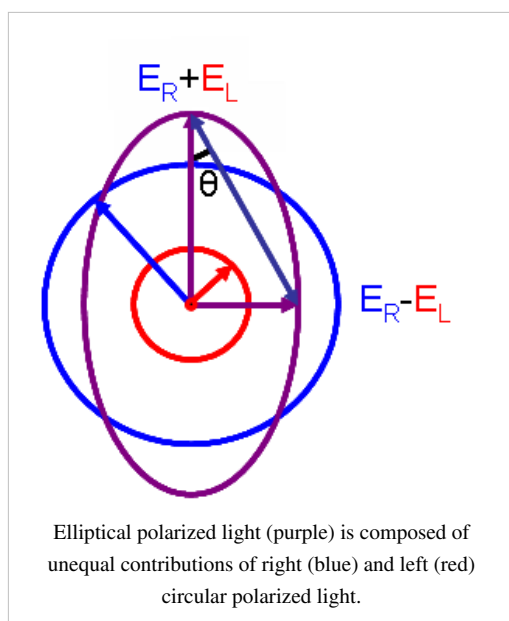
is the molar circular dichroism. This intrinsic property is what is usually meant by the circular dichroism of the substance. Since $\Delta\epsilon$ is a function of wavelength, a molar circular dichroism value ($\Delta\epsilon$) must specify the wavelength at which it is valid.

Extrinsic effects on circular dichroism

In many practical applications of circular dichroism (CD), as discussed below, the measured CD is not simply an intrinsic property of the molecule, but rather depends on the molecular conformation. In such a case the CD may also be a function of temperature, concentration, and the chemical environment, including solvents. In this case the reported CD value must also specify these other relevant factors in order to be meaningful.

Molar ellipticity

Although ΔA is usually measured, for historical reasons most measurements are reported in degrees of ellipticity. Molar circular dichroism and molar ellipticity, $[\theta]$, are readily interconverted by the equation:



$$[\theta] = 3298.2 \Delta \varepsilon.$$

This relationship is derived by defining the ellipticity of the polarization as:

$$\tan \theta = \frac{E_R - E_L}{E_R + E_L}$$

where

E_R and E_L are the magnitudes of the electric field vectors of the right-circularly and left-circularly polarized light, respectively.

When E_R equals E_L (when there is no difference in the absorbance of right- and left-circular polarized light), θ is 0° and the light is linearly polarized. When either E_R or E_L is equal to zero (when there is complete absorbance of the circular polarized light in one direction), θ is 45° and the light is circularly polarized.

Generally, the circular dichroism effect is small, so $\tan \theta$ is small and can be approximated as θ in radians. Since the intensity or irradiance, I , of light is proportional to the square of the electric-field vector, the ellipticity becomes:

$$\theta(\text{radians}) = \frac{(I_R^{1/2} - I_L^{1/2})}{(I_R^{1/2} + I_L^{1/2})}$$

Then by substituting for I using Beer's law in natural logarithm form:

$$I = I_0 e^{-A \ln 10}$$

The ellipticity can now be written as:

$$\theta(\text{radians}) = \frac{(e^{\frac{-A_R}{2} \ln 10} - e^{\frac{-A_L}{2} \ln 10})}{(e^{\frac{-A_R}{2} \ln 10} + e^{\frac{-A_L}{2} \ln 10})} = \frac{e^{\Delta A \frac{\ln 10}{2}} - 1}{e^{\Delta A \frac{\ln 10}{2}} + 1}$$

Since $\Delta A \ll 1$, this expression can be approximated by expanding the exponentials in a Taylor series to first-order and then discarding terms of ΔA in comparison with unity and converting from radians to degrees:

$$\theta(\text{degrees}) = \Delta A \left(\frac{\ln 10}{4} \right) \left(\frac{180}{\pi} \right)$$

The linear dependence of solute concentration and pathlength is removed by defining molar ellipticity as,

$$[\theta] = \frac{100\theta}{Cl}$$

Then combining the last two expression with Beer's law, molar ellipticity becomes:

$$[\theta] = 100 \Delta\epsilon \left(\frac{\ln 10}{4} \right) \left(\frac{180}{\pi} \right) = 3298.2 \Delta\epsilon$$

Mean residue ellipticity

Methods for estimating secondary structure in polymers, proteins and polypeptides in particular, often require that the measured molar ellipticity spectrum be converted to a normalized value, specifically a value independent of the polymer length. Mean residue ellipticity is used for this purpose; it is simply the measured molar ellipticity of the molecule divided by the number of monomer units (residues) in the molecule.

Application to biological molecules

In general, this phenomenon will be exhibited in absorption bands of any optically active molecule. As a consequence, circular dichroism is exhibited by biological molecules, because of their dextrorotary and levorotary components. Even more important is that a secondary structure will also impart a distinct CD to its respective molecules. Therefore, the alpha helix of proteins and the double helix of nucleic acids have CD spectral signatures representative of their structures.

CD is closely related to the optical rotatory dispersion (ORD) technique, and is generally considered to be more advanced. CD is measured in or near the absorption bands of the molecule of interest, while ORD can be measured far from these bands. CD's advantage is apparent in the data analysis. Structural elements are more clearly distinguished since their recorded bands do not overlap extensively at particular wavelengths as they do in ORD. In principle these two spectral measurements can be interconverted through an integral transform (Kramers–Kronig relation), if all the absorptions are included in the measurements.

The far-UV (ultraviolet) CD spectrum of proteins can reveal important characteristics of their secondary structure. CD spectra can be readily used to estimate the fraction of a molecule that is in the alpha-helix conformation, the beta-sheet conformation, the beta-turn conformation, or some other (e.g. random coil) conformation.^{[11] [12]} These fractional assignments place important constraints on the possible secondary conformations that the protein can be in. CD cannot, in general, say where the alpha helices that are detected are located within the molecule or even completely predict how many there are. Despite this, CD is a valuable tool, especially for showing changes in conformation. It can, for instance, be used to study how the secondary structure of a molecule changes as a function of temperature or of the concentration of denaturing agents, e.g. Guanidinium hydrochloride or urea. In this way it can reveal important thermodynamic information about the molecule (such as the enthalpy and Gibbs free energy of denaturation) that cannot otherwise be easily obtained. Anyone attempting to study a protein will find CD a valuable tool for verifying that the protein is in its native conformation before undertaking extensive and/or expensive experiments with it. Also, there are a number of other uses for CD spectroscopy in protein chemistry not related to alpha-helix fraction estimation.

The near-UV CD spectrum (>250 nm) of proteins provides information on the tertiary structure. The signals obtained in the 250–300 nm region are due to the absorption, dipole orientation and the nature of the surrounding environment of the phenylalanine, tyrosine, cysteine (or S–S disulfide bridges) and tryptophan amino acids. Unlike in far-UV CD, the near-UV CD spectrum cannot be assigned to any particular 3D structure. Rather, near-UV CD spectra provide structural information on the nature of the prosthetic groups in proteins, e.g., the heme groups in hemoglobin and cytochrome c.

Visible CD spectroscopy is a very powerful technique to study metal–protein interactions and can resolve individual d–d electronic transitions as separate bands. CD spectra in the visible light region are only produced when a metal ion is in a chiral environment, thus, free metal ions in solution are not detected. This has the advantage of only observing the protein-bound metal, so pH dependence and stoichiometries are readily obtained. Optical activity in transition metal ion complexes have been attributed to configurational, conformational and the vicinal effects. Klewpatinond and Viles (2007) have produced a set of empirical rules for predicting the appearance of visible CD

spectra for Cu^{2+} and Ni^{2+} square-planar complexes involving histidine and main-chain coordination.

CD gives less specific structural information than X-ray crystallography and protein NMR spectroscopy, for example, which both give atomic resolution data. However, CD spectroscopy is a quick method that does not require large amounts of proteins or extensive data processing. Thus CD can be used to survey a large number of solvent conditions, varying temperature, pH, salinity, and the presence of various cofactors.

CD spectroscopy is usually used to study proteins in solution, and thus it complements methods that study the solid state. This is also a limitation, in that many proteins are embedded in membranes in their native state, and solutions containing membrane structures are often strongly scattering. CD is sometimes measured in thin films.

Experimental limitations

CD has also been studied in carbohydrates, but with limited success due to the experimental difficulties associated with measurement of CD spectra in the vacuum ultraviolet (VUV) region of the spectrum (100-200 nm), where the corresponding CD bands of unsubstituted carbohydrates lie. Substituted carbohydrates with bands above the VUV region have been successfully measured.

Measurement of CD is also complicated by the fact that typical aqueous buffer systems often absorb in the range where structural features exhibit differential absorption of circularly polarized light. Phosphate, sulfate, carbonate, and acetate buffers are generally incompatible with CD unless made extremely dilute e.g. in the 10-50 mM range. The TRIS buffer system should be completely avoided when performing far-UV CD. Borate and Onium compounds are often used to establish the appropriate pH range for CD experiments. Some experimenters have substituted fluoride for chloride ion because fluoride absorbs less in the far UV, and some have worked in pure water. Another, almost universal, technique is to minimize solvent absorption by using shorter path length cells when working in the far UV, 0.1 mm path lengths are not uncommon in this work.

In addition to measuring in aqueous systems, CD, particularly far-UV CD, can be measured in organic solvents e.g. ethanol, methanol, trifluoroethanol (TFE). The latter has the advantage to induce structure formation of proteins, inducing beta-sheets in some and alpha helices in others, which they would not show under normal aqueous conditions. Most common organic solvents such as acetonitrile, THF, chloroform, dichloromethane are however, incompatible with far-UV CD.

It may be of interest to note that the protein CD spectra used in secondary structure estimation are related to the π to π^* orbital absorptions of the amide bonds linking the amino acids. These absorption bands lie partly in the *so-called* vacuum ultraviolet (wavelengths less than about 200 nm). The wavelength region of interest is actually inaccessible in **air** because of the strong absorption of light by oxygen at these wavelengths. In practice these spectra are measured not in vacuum but in an oxygen-free instrument (filled with pure nitrogen gas).

Once oxygen has been eliminated, perhaps the second most important technical factor in working below 200 nm is to design the rest of the optical system to have low losses in this region. Critical in this regard is the use of aluminized mirrors whose coatings have been optimized for low loss in this region of the spectrum.

The usual light source in these instruments is a high pressure, short-arc xenon lamp. Ordinary xenon arc lamps are unsuitable for use in the low UV. Instead, specially constructed lamps with envelopes made from high-purity synthetic fused silica must be used.

Light from synchrotron sources has a much higher flux at short wavelengths, and has been used to record CD down to 160 nm. Recently the CD spectrometer at the electron storage ring facility ISA at the University of Aarhus in Denmark was used to record solid state CD spectra down to 120 nm.

At the quantum mechanical level, the information content of circular dichroism and optical rotation are identical.

See also

- Circular polarization in nature
- Dichroism
- Linear dichroism
- Magnetic circular dichroism
- Optical activity
- Optical isomerism
- Optical rotation
- Optical rotatory dispersion

Further reading

1. Alison Rodger and Bengt Nordén, *Circular Dichroism and Linear Dichroism* ^[13] (1997) Oxford University Press, Oxford, UK. ISBN 019855897X.
2. Fasman, G.D., *Circular Dichroism and the Conformational Analysis of Biomolecules* (1996) Plenum Press, New York.
3. Hecht, E., *Optics* 3rd Edition (1998) Addison Wesley Longman, Massachusetts.
4. Klewpatinond, M. and Viles, J.H. (2007) Empirical rules for rationalising visible circular dichroism of Cu²⁺ and Ni²⁺ histidine complexes: Applications to the prion protein. *FEBS Letters* 581, 1430-1434.

External links

- Circular Dichroism explained ^[14]
- Circular Dichroism at UMDNJ ^[15] - a good site for information on structure estimation software
- Electromagnetic waves ^[16] - Animated electromagnetic waves. The Emanim program is a teaching resource for helping students understand the nature of electromagnetic waves and their interaction with birefringent and dichroic samples
- An Introduction to Circular Dichroism Spectroscopy ^[17] - a very good tutorial on circular dichroism spectroscopy

References

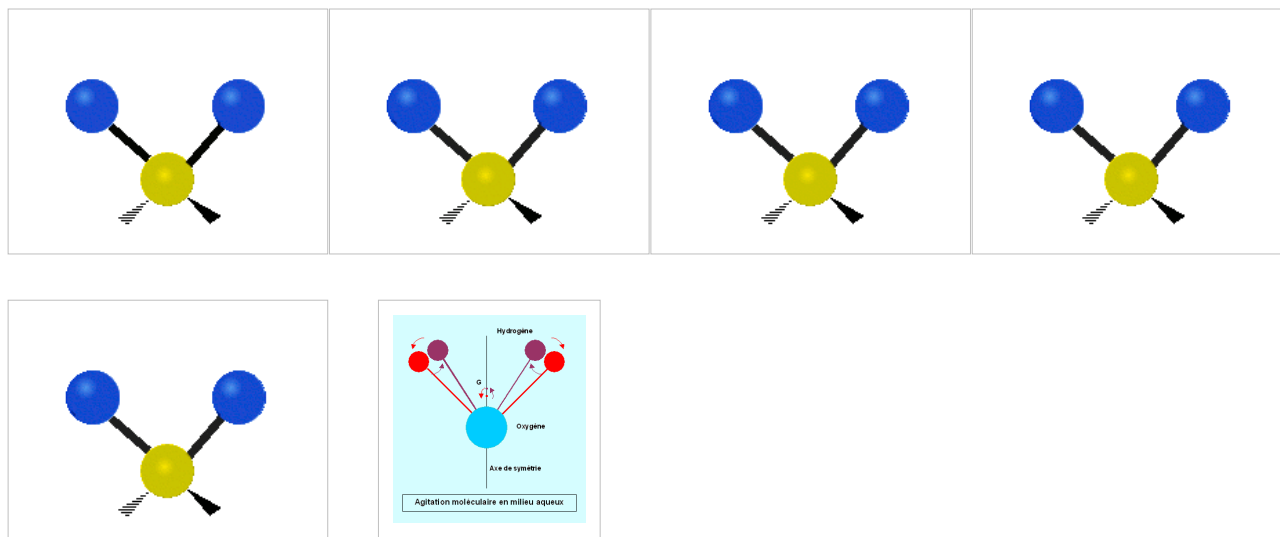
- [1] G. D. Fasman (1996). Plenum Press. p. 3.
- [2] P. Atkins and J. de Paula (2005). *Elements of Physical Chemistry, 4th ed.*. Oxford University Press.
- [3] E. I. Solomon and A. B. P. Lever (2006). **1**. Wiley. p. 78.
- [4] R. W. Woody (1994). K. Nakanishi, N. Berova, R. W. Woody. ed. VCH Publishers, Inc.. p. 473.
- [5] Solomon, Neidig; A. T. Wecksler, G. Schenk, and T. R. Holman (2007). "Kinetic and Spectroscopic Studies of N694C Lipoxygenase: A Probe of the Substrate Activation Mechanism of a Nonheme Ferric Enzyme". *JACS* **129**: 7531–7537.
- [6] E. I. Solomon and A. B. P. Lever (2006). **1**. Wiley. p. 78.
- [7] K. Nakanishi, N. Berova, R. W. Woody, ed (1994). VCH Publishers, Inc..
- [8] A. Rodger and B. Norden (1997). *Circular Dichroism and Linear Dichroism*. Oxford University Press.
- [9] E. I. Solomon and A. B. P. Lever (2006). **1**. Wiley. p. 78.
- [10] Woody, R. W. 1994
- [11] Whitmore L, Wallace BA (2008). "Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases". *Biopolymers* **89** (5): 392–400. doi:10.1002/bip.20853. PMID 17896349.
- [12] Greenfield NJ (2006). "Using circular dichroism spectra to estimate protein secondary structure". *Nature protocols* **1** (6): 2876–90. doi:10.1038/nprot.2006.202. PMID 17406547.
- [13] <http://books.google.co.uk/books?hl=en&id=ThEKGC99hJcC>
- [14] http://www.ap-lab.com/circular_dichroism.htm
- [15] <http://www2.umdj.edu/cdrwjweb/index.htm#software>
- [16] <http://www.enzim.hu/~szia/cddemo/demo1.htm>
- [17] <http://www.photophysics.com/circulardichroism.php>

Vibrational circular dichroism

Vibrational circular dichroism (VCD) is a spectroscopic technique which detects differences in attenuation of left and right circularly polarized light passing through a sample. It is basically circular dichroism spectroscopy in the infrared and near infrared ranges^[1].

Because VCD is sensitive to the mutual orientation of distinct groups in a molecule, it provides three-dimensional structural information. Thus, it is a powerful technique as VCD spectra of enantiomers can be simulated using *ab initio* calculations, thereby allowing the identification of absolute configurations of small molecules in solution from VCD spectra. Among such quantum computations of VCD spectra resulting from the chiral properties of small organic molecules are those based on density functional theory (DFT) and gauge-invariant atomic orbitals (GIAO). As a simple example of the experimental results that were obtained by VCD are the spectral data obtained within the carbon-hydrogen (C-H) stretching region of 21 amino acids in heavy water solutions. Measurements of vibrational optical activity (VOA) have thus numerous applications, not only for small molecules, but also for large and complex biopolymers such as muscle proteins (myosin, for example) and DNA.

Vibrational modes



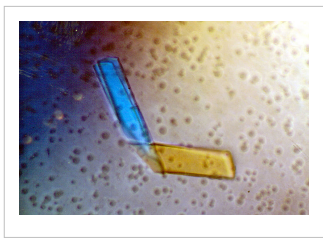
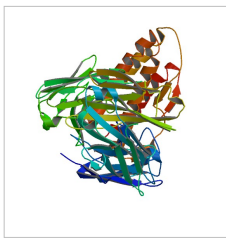
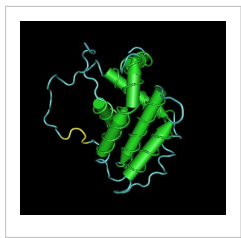
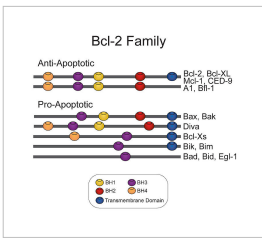
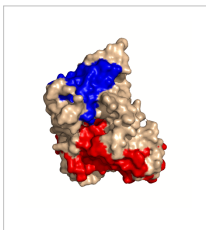
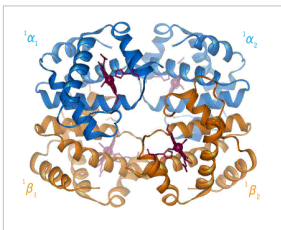
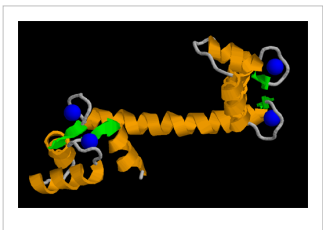
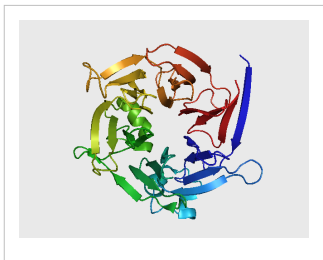
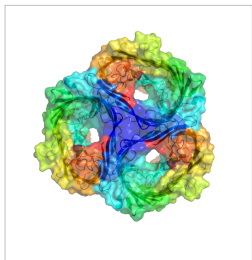
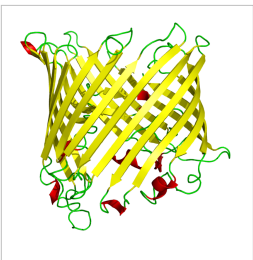
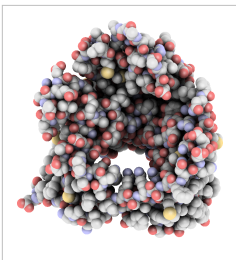
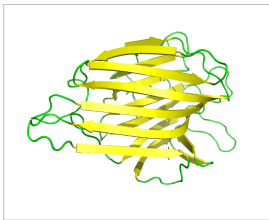
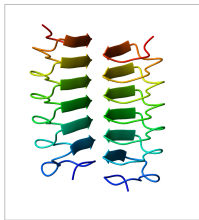
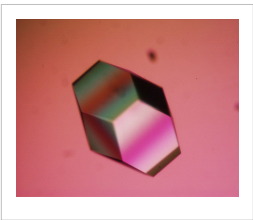
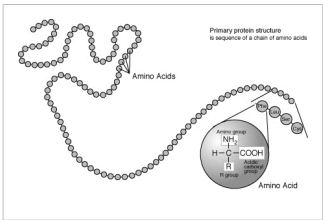
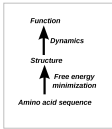
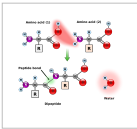
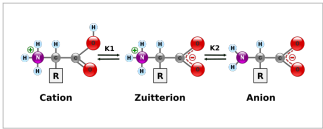
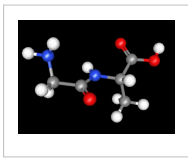
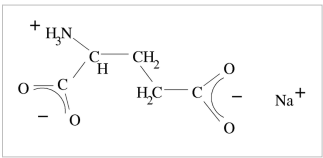
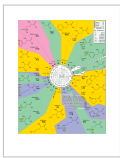
Theory of VCD

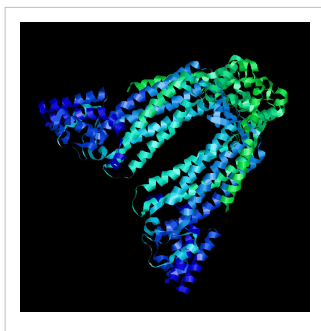
While the fundamental quantity associated with the infrared absorption is the dipole strength, the differential absorption is proportional also to the rotational strength, a quantity which depends on both the electric and magnetic dipole transition moments. Sensitivity of the handedness of a molecule toward circularly polarized light results from the form of the rotational strength.

VCD of peptides and proteins

Extensive VCD studies have been reported for both polypeptides and several proteins in solution^{[2] [3] [4]}; several recent reviews were also compiled^{[5] [6] [7] [8]}. An extensive but not comprehensive VCD publications list is also provided in the "References" section. The published reports over the last 22 years have established VCD as a powerful technique with improved results over those previously obtained by visible/UV circular dichroism (CD) or optical rotatory dispersion (ORD) for proteins and nucleic acids.

Amino acid and polypeptide structures



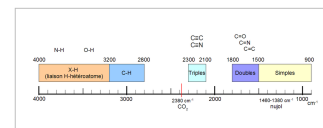
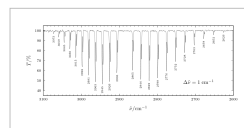
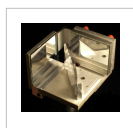
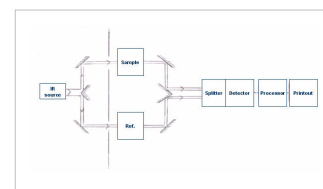
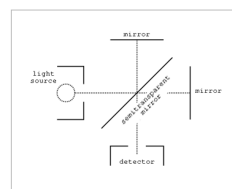
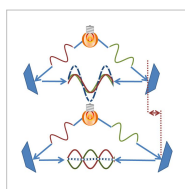
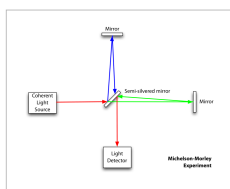


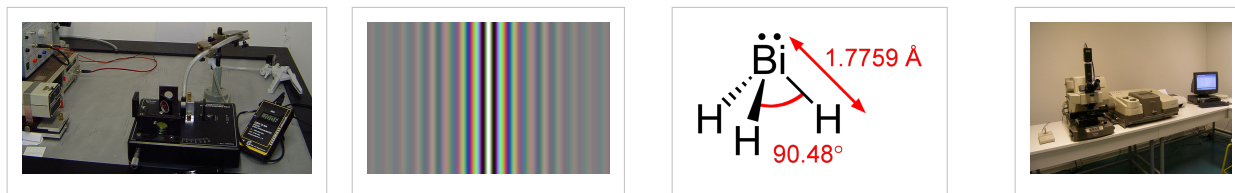
VCD of nucleic acids

VCD spectra of nucleotides, synthetic polynucleotides and several nucleic acids, including DNA, have been reported and assigned in terms of the type and number of helices present in A-, B-, and Z- DNA.

VCD Instrumentation

For biopolymers such as proteins and nucleic acids, the difference in absorbance between the levo- and dextro-configurations is five orders of magnitude smaller than the corresponding (unpolarized) absorbance. Therefore, VCD of biopolymers requires the use of very sensitive, specially built instrumentation as well as time-averaging over relatively long intervals of time even with such sensitive VCD spectrometers. Most CD instruments produce left- and right- circularly polarized light which is then either sine-wave or square-wave modulated, with subsequent phase-sensitive detection and lock-in amplification of the detected signal. In the case of FT-VCD, a photo-elastic modulator (PEM) is employed in conjunction with an FT-IR interferometer set-up. An example is that of a Bomem model MB-100 FT-IR interferometer equipped with additional polarizing optics/ accessories needed for recording VCD spectra. A parallel beam emerges through a side port of the interferometer which passes first through a wire grid linear polarizer and then through an octagonal-shaped ZnSe crystal PEM which modulates the polarized beam at a fixed, lower frequency such as 37.5 kHz. A mechanically stressed crystal such as ZnSe exhibits birefringence when stressed by an adjacent piezoelectric transducer. The linear polarizer is positioned close to, and at 45 degrees, with respect to the ZnSe crystal axis. The polarized radiation focused onto the detector is doubly modulated, both by the PEM and by the interferometer setup. A very low noise detector, such as MCT (HgCdTe), is also selected for the VCD signal phase-sensitive detection. Quasi-complete commercial FT-VCD instruments are also available from a few manufacturers but these are quite expensive and also have to be still considered as being at the prototype stage. To prevent detector saturation an appropriate, long wave pass filter is placed before the very low noise MCT detector, which allows only radiation below 1750 cm^{-1} to reach the MCT detector; the latter however measures radiation only down to 750 cm^{-1} . FT-VCD spectra accumulation of the selected sample solution is then carried out, digitized and stored by an in-line computer. Published reviews that compare various VCD methods are also available.^{[9] [10]}





Magnetic VCD

VCD spectra have also been reported in the presence of an applied external magnetic field^[11]. This method can enhance the VCD spectral resolution for small molecules^{[12] [13] [14] [15] [16]}.

Raman optical activity (ROA)

ROA is a technique complementary to VCD especially useful in the 50—1600 cm^{-1} spectral region; it is considered as the technique of choice for determining optical activity for photon energies less than 600 cm^{-1} .

References

Peptides and proteins

- Huang R, Wu L, McElheny D, Bour P, Roy A, Keiderling TA. Cross-Strand Coupling and Site-Specific Unfolding Thermodynamics of a Trpzip beta-Hairpin Peptide Using (13)C Isotopic Labeling and IR Spectroscopy. *The journal of physical chemistry. B*. 2009 Apr;113(16):5661-74.
- "Vibrational Circular Dichroism of Poly alpha-Benzyl-L-Glutamate," R. D. Singh, and T. A. Keiderling, *Biopolymers*, 20, 237-40 (1981).
- "Vibrational Circular Dichroism of Polypeptides II. Solution Amide II and Deuteration Results," A. C. Sen and T. A. Keiderling, *Biopolymers*, 23, 1519-32 (1984).
- "Vibrational Circular Dichroism of Polypeptides III. Film Studies of Several alpha-Helical and β -Sheet Polypeptides," A. C. Sen and T. A. Keiderling, *Biopolymers*, 23, 1533-46 (1984).
- "Vibrational Circular Dichroism of Polypeptides IV. Film Studies of L-Alanine Homo Oligopeptides," U. Narayanan, T. A. Keiderling, G. M. Bonora, and C. Toniolo, *Biopolymers* 24, 1257-63 (1985).
- "Vibrational Circular Dichroism of Polypeptides, T. A. Keiderling, S. C. Yasui, A. C. Sen, C. Toniolo, G. M. Bonora, in *Peptides Structure and Function, Proceedings of the 9th American Peptide Symposium*," ed. C. M. Deber, K. Kopple, V. Hruby; Pierce Chemical: Rockford, IL; 167-172 (1985).
- "Vibrational Circular Dichroism of Polypeptides V. A Study of 310 Helical-Octapeptides" S. C. Yasui, T. A. Keiderling, G. M. Bonora, C. Toniolo, *Biopolymers* 25, 79-89 (1986).
- "Vibrational Circular Dichroism of Polypeptides VI. Polytyrosine alpha-helical and Random Coil Results," S. C. Yasui and T. A. Keiderling, *Biopolymers* 25, 5-15 (1986).
- "Vibrational Circular Dichroism of Polypeptides VII. Film and Solution Studies of alpha-forming Homo-Oligopeptides," U. Narayanan, T. A. Keiderling, G. M. Bonora, C. Toniolo, *Journal of the American Chemical Society*, 108, 2431-2437 (1986).
- "Vibrational Circular Dichroism of Polypeptides VIII. Poly Lysine Conformations as a Function of pH in Aqueous Solution," S. C. Yasui, T. A. Keiderling, *Journal of the American Chemical Society*, 108, 5576-5581 (1986).
- "Vibrational Circular Dichroism of Polypeptides IX. A Study of Chain Length Dependence for 310-Helix Formation in Solution." S. C. Yasui, T. A. Keiderling, F. Formaggio, G. M. Bonora, C. Toniolo, *Journal of the American Chemical Society* 108, 4988-4993 (1986).
- "Vibrational Circular Dichroism of Biopolymers." T. A. Keiderling, *Nature*, 322, 851-852 (1986).

- "Vibrational Circular Dichroism of Polypeptides X. A Study of alpha-Helical Oligopeptides in Solution." S. C. Yasui, T. A. Keiderling, R. Katachai, *Biopolymers* 26, 1407-1412 (1987).
- "Vibrational Circular Dichroism of Polypeptides XI. Conformation of Poly(L-Lysine(Z)-L-Lysine(Z)-L-1-Pyrenylalanine) and Poly(L-Lysine(Z)-L-Lysine(Z)-L-1-Naphthylalanine) in Solution" S. C. Yasui, T. A. Keiderling, and M. Sisido, *Macromolecules* 20, 2 403-2406 (1987).
- "Vibrational Circular Dichroism of Biopolymers" T. A. Keiderling, S. C. Yasui, A. C. Sen, U. Narayanan, A. Annamalai, P. Malon, R. Kobrinskaya, L. Yang, in "F.E.C.S. Second International Conference on Circular Dichroism, Conference Proceedings," ed. M. Kajtar, L. Eötvös Univ., Budapest, 1987, p. 155-161.
- "Vibrational Circular Dichroism of Poly-L-Proline and Other Helical Poly-peptides," R. Kobrinskaya, S. C. Yasui, T. A. Keiderling, in "Peptides: Chemistry and Biology, Proceedings of the 10th American Peptide Symposium," ed. G. R. Marshall, ESCOM, Leiden, 1988, p. 65-67.
- "Vibrational Circular Dichroism of Polypeptides with Aromatic Side Chains," S. C. Yasui, T. A. Keiderling, in "Peptides: Chemistry and Biology, Proceedings of the 10th American Peptide Symposium," ed. G. R. Marshall, ESCOM, Leiden, 1988, p. 90-92.
- "Vibrational Circular Dichroism of Polypeptides XII. Re-evaluation of the Fourier Transform Vibrational Circular Dichroism of Poly-gamma-Benzyl-L-Glutamate," P. Malon, R. Kobrinskaya, T. A. Keiderling, *Biopolymers* 27, 733-746 (1988).
- "Vibrational Circular Dichroism of Biopolymers," T. A. Keiderling, S. C. Yasui, U. Narayanan, A. Annamalai, P. Malon, R. Kobrinskaya, L. Yang, in *Spectroscopy of Biological Molecules New Advances* ed. E. D. Schmid, F. W. Schneider, F. Siebert, p. 73-76 (1988).
- "Vibrational Circular Dichroism of Polypeptides and Proteins," S. C. Yasui, T. A. Keiderling, *Mikrochimica Acta*, II, 325-327, (1988).
- "(1R,7R)-7-Methyl-6,9,-Diazatricyclo[6,3,0,01,6]Tridecane-5,10-Dione, A Tricyclic Spirodilactam Containing Non-planar Amide Groups: Synthesis, NMR, Crystal Structure, Absolute Configuration, Electronic and Vibrational Circular Dichroism" P. Malon, C. L. Barnes, M. Budesinsky, R. K. Dukor, D. van der Helm, T. A. Keiderling, Z. Koblicova, F. Pavlikova, M. Tichy, K. Blaha, *Collections of Czechoslovak Chemical Communications* 53, 2447-2472 (1988).
- "Vibrational Circular Dichroism of Poly Glutamic Acid" R. K. Dukor, T. A. Keiderling, in *Peptides 1988* (ed. G. Jung, E. Bayer) Walter de Gruyter, Berlin (1989) pp 519-521.
- "Biopolymer Conformational Studies with Vibrational Circular Dichroism" T. A. Keiderling, S. C. Yasui, P. Pancoska, R. K. Dukor, L. Yang, SPIE Proceeding 1057, ("Biomolecular Spectroscopy," ed. H. H. Mantsch, R. R. Birge) 7-14 (1989).
- "Vibrational Circular Dichroism. Comparison of Techniques and Practical Considerations" T. A. Keiderling, in "Practical Fourier Transform Infrared Spectroscopy. Industrial and Laboratory Chemical Analysis," ed. J. R. Ferraro, K. Krishnan (Academic Press, San Diego, 1990) p. 203-284.
- "Vibrational Circular Dichroism Study of Unblocked Proline Oligomers," R. K. Dukor, T. A. Keiderling, V. Gut, *International Journal of Peptide and Protein Research*, 38, 198-203 (1991).
- "Reassessment of the Random Coil Conformation. Vibrational CD Study of Proline Oligopeptides and Related Polypeptides" R. K. Dukor and T. A. Keiderling, *Biopolymers* 31 1747-1761 (1991).
- "Vibrational CD of the Amide II band in Some Model Polypeptides and Proteins" V. P. Gupta, T. A. Keiderling, *Biopolymers* 32 239-248 (1992).
- "Vibrational Circular Dichroism of Proteins Polysaccharides and Nucleic Acids" T. A. Keiderling, Chapter 8 in *Physical Chemistry of Food Processes, Vol. 2 Advanced Techniques, Structures and Applications.*, eds. I.C. Baianu, H. Pessen, T. Kumosinski, Van Norstrand—Reinhold, New York (1993), pp 307-337.
- "Structural Studies of Biological Macromolecules using Vibrational Circular Dichroism" T. A. Keiderling, P. Pancoska, Chapter 6 in *Advances in Spectroscopy Vol. 21, Biomolecular Spectroscopy Part B* eds. R. E. Hester, R. J. H. Clarke, John Wiley Chichester (1993) pp 267-315.

- "Ab Initio Simulations of the Vibrational Circular Dichroism of Coupled Peptides" P. Bour and T. A. Keiderling, *Journal of the American Chemical Society* 115 9602-9607 (1993).
- "Ab initio Simulations of Coupled Peptide Vibrational Circular Dichroism" P. Bour, T. A. Keiderling in "Fifth International Conference on The Spectroscopy of Biological Molecules" Th. Theophanides, J. Anastassopoulou, N. Fotopoulos (Eds), Kluwer Academic Publ., Dordrecht, 1993, p. 29-30.
- "Vibrational Circular Dichroism Spectroscopy of Peptides and Proteins" T. A. Keiderling, in "Circular Dichroism Interpretations and Applications," K. Nakanishi, N. Berova, R. Woody, Eds., VCH Publishers, New York, (1994) pp 497-521.
- "Conformational Study of Sequential Lys-Leu Based Polymers and Oligomers using Vibrational and Electronic Circular Dichroism Spectra" V. Baumruk, D. Huo, R. K. Dukor, T. A. Keiderling, D. LeLeivre and A. Brack *Biopolymers* 34, 1115-1121 (1994).
- "Vibrational Optical Activity of Oligopeptides" T. B. Freedman, L. A. Nafie, T. A. Keiderling *Biopolymers* (Peptide Science) 37 (ed. C. Toniolo) 265-279 (1995).
- "Characterization of β -bend ribbon spiral forming peptides using electronic and vibrational circular dichroism" G. Yoder, T. A. Keiderling, F. Formaggio, M. Crisma, C. Toniolo *Biopolymers* 35, 103-111 (1995).
- "Vibrational Circular Dichroism as a Tool for Determination of Peptide Secondary Structure" P. Bour, T. A. Keiderling, P. Malon, in "Peptides 1994 (Proceedings of the 23rd European Peptide Symposium, 1994," (H.L.S. Maia, ed.), Escom, Leiden 1995, p. 517-518.
- "Helical Screw Sense of homo-oligopeptides of C-alpha-methylated alpha-amino acids as Determined with Vibrational Circular Dichroism." G. Yoder, T. A. Keiderling, M. Crisma, F. Formaggio, C. Toniolo, J. Kamphuis, *Tetrahedron Asymmetry* 6, 687 -690 (1995).
- "Conformational Study of Linear Alternating and Mixed D- and L-Proline Oligomers Using Electronic and Vibrational CD and Fourier Transform IR." W. M. Stille, R. K. Dukor, G. Yoder, T. A. Keiderling *Biopolymers* 36, 623-631 (1995).
- Review: "Vibrational Circular Dichroism Applications to Conformational Analysis of Biomolecules" T. A. Keiderling in *Circular Dichroism and the Conformational Analysis of Biomolecules* ed. G. D. Fasman, Plenum, New York (1996) p. 555-585.
- "Mutarotation studies of Poly L-Proline using FT-IR, Electronic and Vibrational Circular Dichroism" R. K. Dukor, T. A. Keiderling, *Biospectroscopy* 2, 83-100 (1996).
- "Vibrational Circular Dichroism Applications in Proteins and Peptides" T. A. Keiderling, *Proceedings of the NATO ASI in Biomolecular Structure and Dynamics, Loutraki Greece, May 1996*, Ed. G. Vergoten (delayed second volume to 1998).
- "Transfer of Molecular Property Tensors in Cartesian Coordinates: A new algorithm for simulation of vibrational spectra" Petr Bour, Jana Sopkova, Lucie Bednarova, Petr Malon, T. A. Keiderling, *Journal of Computational Chemistry* 18, 646-659 (1997).
- "Vibrational Circular Dichroism Characterization of Alanine-Rich Peptides." Gorm Yoder and Timothy A. Keiderling, "Spectroscopy of Biological Molecules: Modern Trends," Ed. P. Carmona, R. Navarro, A. Hernanz, Kluwer Acad. Pub., Netherlands (1997) p. 27-28.
- "Ionic strength effect on the thermal unfolding of alpha-spectrin peptides." D. Lusitani, N. Menhart, T.A. Keiderling and L. W. M. Fung. *Biochemistry* 37(1998)16546-16554.
- "In search of the earliest events of hCgb folding: structural studies of the 60-87 peptide fragment" S. Sherman, L. Kirnarsky, O. Prakash, H. M. Rogers, R.A.G.D. Silva, T.A. Keiderling, D. Smith, A.M. Hanly, F. Perini, and R.W. Ruddon, *American Peptide Symposium Proceedings*, 1997.
- "Cold Denaturation Studies of (LKELPKEL)_n Peptide Using Vibrational Circular Dichroism and FT-IR". R. A. G. D. Silva, Vladimir Baumruk, Petr Pancoska, T. A. Keiderling, Eric Lacassie, and Yves Trudelle, *American Peptide Symposium Proceedings*, 1997.

- "Simulations of oligopeptide vibrational CD. Effects of isotopic labeling." Petr Bour, Jan Kubelka, T. A. Keiderling *Biopolymers* 53, 380-395 (2000).
- "Site specific conformational determination in thermal unfolding studies of helical peptides using vibrational circular dichroism with isotopic substitution" R. A. G. D. Silva, Jan Kubelka, Petr Bour, Sean M. Decatur, Timothy A. Keiderling, *Proceedings of the National Academy of Sciences* (PNAS:USA) 97, 8318-8323 (2000).
- "Folding studies on the human chorionic gonadotropin b -subunit using optical spectroscopy of peptide fragments" R. A. G. D. Silva, S. A. Sherman, F. Perini, E. Bedows, T. A. Keiderling, *Journal of the American Chemical Society*, 122, 8623-8630 (2000).
- "Peptide and Protein Conformational Studies with Vibrational Circular Dichroism and Related Spectroscopies", Timothy A. Keiderling, (Revised and Expanded Chapter) In *Circular Dichroism: Principles and Applications*, 2nd Edition. (Eds. K. Nakanishi, N. Berova and R. A. Woody, John Wiley & Sons, New York (2000) p. 621-666.
- "Conformation studies with Optical Spectroscopy of peptides taken from hairpin sequences in the Human Chorionic Gonadotropin " R. A. G. D. Silva, S. A. Sherman, E. Bedows, T. A. Keiderling, *Peptides for the New Millenium, Proceedings of the 16th American Peptide Symposium*, (June, 1999 Minneapolis, MN) Ed. G. B. Fields, J. P. Tam, G. Barany, Kluwer Acad. Pub., Dordrecht, (2000) p. 325-326.
- "Analysis of Local Conformation within Helical Peptides via Isotope-Edited Vibrational Spectroscopy." S. M. Decatur, T. A. Keiderling, R. A. G. D. Silva, and P. Bour, *Peptides for the New Millenium, Proceedings of the 16th American Peptide Symposium*, (June, 1999 Minneapolis, MN) Ed. G. B. Fields, J. P. Tam, G. Barany, Kluwer Acad. Pub., Dordrecht, (2000) p. 414-416.
- "The anomalous infrared amide I intensity distribution in C-13 isotopically labeled peptide beta-sheets comes from extended, multiple stranded structures. An *Ab Initio* study." Jan Kubelka and T. A. Keiderling, *Journal of the American Chemical Society*. 123, 6142-6150 (2001).
- "Vibrational Circular Dichroism of Peptides and Proteins: Survey of Techniques, Qualitative and Quantitative Analyses, and Applications" Timothy A. Keiderling, Chapter in *Infrared and Raman Spectroscopy of Biological Materials*, Ed. Bing Yan and H.-U. Gremlich, Marcel Dekker, New York (2001) p. 55-100.
- "Chirality in peptide vibrations. Ab initio computational studies of length, solvation, hydrogen bond, dipole coupling and isotope effects on vibrational CD. " Jan Kubelka, Petr Bour, R. A. Gangani D. Silva, Sean M. Decatur, Timothy A. Keiderling, *ACS Symposium Series* 810, ["Chirality: Physical Chemistry," (Ed. Janice Hicks) American Chemical Society, Washington, DC] (2002), pp. 50–64.
- "Spectroscopic Characterization of Selected b-Sheet Hairpin Models", J. Hilario, J. Kubelka, F. A. Syud, S. H. Gellman, and T. A. Keiderling. *Biopolymers (Biospectroscopy)* 67: 233-236 (2002)
- " Discrimination between peptide 3_{10} - and alpha-helices. Theoretical analysis of the impact of alpha-methyl substitution on experimental spectra " Jan Kubelka, R. A. Gangani D. Silva, and T. A. Keiderling, *Journal of the American Chemical Society*, 124, 5325-5332 (2002).
- "*Ab Initio* Quantum Mechanical Models of Peptide Helices and their Vibrational Spectra" Petr Bour, Jan Kubelka and T. A. Keiderling, *Biopolymers* 65, 45-59 (2002).
- "Discriminating 3_{10} - from alpha-helices. Vibrational and electronic CD and IR Absorption study of related Aib-containing oligopeptides" R. A. Gangani D. Silva, Sritana Yasui, Jan Kubelka, Fernando Formaggio, Marco Crisma, Claudio Toniolo, and Timothy A. Keiderling, *Biopolymers* 65, 229-243 (2002).
- "Spectroscopic characterization of Unfolded peptides and proteins studied with infrared absorption and vibrational circular dichroism spectra" T. A. Keiderling and Qi Xu, *Advances in Protein Chemistry Volume 62*, [Unfolded Proteins, Dedicated to John Edsall, Ed.: George Rose, Academic Press:New York] (2002), pp. 111–161.
- "Protein and Peptide Secondary Structure and Conformational Determination with Vibrational Circular Dichroism " Timothy A. Keiderling, *Current Opinions in Chemical Biology* (Ed. Julie Leary and Mark Arnold) 6, 682-688 (2002).

- Review: Conformational Studies of Peptides with Infrared Techniques. Timothy A. Keiderling and R. A. G. D. Silva, in *Synthesis of Peptides and Peptidomimetics*, Ed. M. Goodman and G. Herrman, Houben-Weyl, Vol 22Eb, Georg Thiem Verlag, New York (2002) pp. 715–738, (written and accepted in 2000).
- "Spectroscopic Studies of Structural Changes in Two beta-Sheet Forming Peptides Show an Ensemble of Structures That Unfold Non-Cooperatively" Serguei V. Kuznetsov, Jovencio Hilario, T. A. Keiderling, Anjum Ansari, *Biochemistry*, 42 :4321-4332, (2003).
- "Optical spectroscopic investigations of model beta-sheet hairpins in aqueous solution" Jovencio Hilario, Jan Kubelka, T. A. Keiderling, *Journal of the American Chemical Society* 125, 7562-7574 (2003).
- "Synthesis and conformational study of homopeptides based on (S)-Bin, a C2-symmetric binaphthyl-derived Caa-disubstituted glycine with only axial chirality" J.-P. Mazaleyrat, K. Wright, A. Gaucher, M. Wakselman, S. Oancea, F. Formaggio, C. Toniolo, V. Setnicka, J. Kapitan, T. A. Keiderling, *Tetrahedron Asymmetry*, 14, 1879-1893 (2003).
- "Empirical modeling of the peptide amide I band IR intensity in water solution," Petr Bour, Timothy A. Keiderling, *Journal of Chemical Physics*, 119, 11253-11262 (2003)
- "The Nature of Vibrational Coupling in Helical Peptides: An Isotope Labeling Study" by R. Huang, J. Kubelka, W. Barber-Armstrong, R. A. G. D Silva, S. M. Decatur, and T. A. Keiderling, *Journal of the American Chemical Society*, 126, 2346-2354 (2004).
- "The Complete Chiro-spectroscopic Signature of the Peptide 3_{10} Helix in Aqueous Solution" Claudio Toniolo, Fernando Formaggio, Sabrina Tognon, Quirinus B. Broxterman, Bernard Kaptein, Rong Huang, Vladimir Setnicka, Timothy A. Keiderling, Iain H. McColl, Lutz Hecht, Laurence D. Barron, *Biopolymers* 75, 32-45 (2004).
- "Induced axial chirality in the biphenyl core for the Ca-tetrasubstituted α -amino acid residue Bip and subsequent propagation of chirality in (Bip)n/Val oligopeptides" J.-P. Mazaleyrat, K. Wright, A. Gaucher, N. Toulemonde, M. Wakselman, S. Oancea, C. Peggion, F. Formaggio, V. Setnicka, T. A. Keiderling, C. Toniolo, *Journal of the American Chemical Society* 126; 12874-12879 (2004).
- *Ab initio* modeling of amide I coupling in anti-parallel β -sheets and the effect of the ^{13}C isotopic labeling on vibrational spectra" Petr Bour, Timothy A. Keiderling, *Journal of Physical Chemistry B*, 109, 5348-5357 (2005)
- Solvent Effects on IR And VCD Spectra of Helical Peptides: Insights from *Ab Initio* Spectral Simulations with Explicit Water" Jan Kubelka and Timothy A. Keiderling, *Journal of Physical Chemistry B* 109, 8231-8243 (2005)
- IR Study of Cross-Strand Coupling in a beta-Hairpin Peptide Using Isotopic Labels., Vladimir Setnicka, Rong Huang, Catherine L. Thomas, Marcus A. Etienne, Jan Kubelka, Robert P. Hammer, Timothy A. Keiderling *Journal of the American Chemical Society* 127, 4992-4993 (2005).
- Vibrational spectral simulation for peptides of mixed secondary structure: Method comparisons with the trpzip model hairpin. Petr Bour and Timothy A. Keiderling, *Journal of Physical Chemistry B* 109, 232687-23697 (2005).
- Isotopically labeled peptides provide site-resolved structural data with infrared spectra. Probing the structural limit of optical spectroscopy, Timothy A. Keiderling, Rong Huang, Jan Kubelka, Petr Bour, Vladimir Setnicka, Robert P. Hammer, Marcus *A. Etienne, R. A. Gangani D. Silva, Sean M. Decatur Collections Symposium Series, 8, 42-49 (2005)—["Biologically Active Peptides" IXth Conference, Prague Czech Republic, April 20-22, 2005.

Nucleic acids and polynucleotides

- "Application of Vibrational Circular Dichroism to Synthetic Polypeptides and Polynucleic Acids" T. A. Keiderling, S. C. Yasui, R. K. Dukor, L. Yang, *Polymer Preprints* 30, 423-424 (1989).
- "Vibrational Circular Dichroism of Polyribonucleic Acids. A Comparative Study in Aqueous Solution." A. Annamalai and T. A. Keiderling, *Journal of the American Chemical Society*, 109, 3125-3132 (1987).
- "Conformational phase transitions (A-B and B-Z) of DNA and models using vibrational circular dichroism" L. Wang, L. Yang, T. A. Keiderling in *Spectroscopy of Biological Molecules.*, eds. R. E. Hester, R. B. Girling, Special Publication 94 Royal Society of Chemistry, Cambridge (1991) p. 137-38.
- "Vibrational Circular Dichroism of Proteins Polysaccharides and Nucleic Acids" T. A. Keiderling, Chapter 8 in *Physical Chemistry of Food Processes, Vol. 2 Advanced Techniques, Structures and Applications* eds. I. C. Baianu, H. Pessen, T. Kumosinski, Van Norstrand—Reinhold, New York (1993) pp. 307–337.
- "Structural Studies of Biological Macromolecules using Vibrational Circular Dichroism" T. A. Keiderling, P. Pancoska, Chapter 6 in *Advances in Spectroscopy Vol. 21, "Biomolecular Spectroscopy Part B"* ed. R. E. Hester, R. J. H. Clarke, John Wiley Chichester (1993) pp 267–315.
- "Detection of Triple Helical Nucleic Acids with Vibrational Circular Dichroism," L. Wang, P. Pancoska, T. A. Keiderling in "Fifth International Conference on The Spectroscopy of Biological Molecules" Th. Theophanides, J. Anastassopoulou, N. Fotopoulos (Eds), Kluwer Academic Publ., Dordrecht, 1993, p. 81-82.
- "Helical Nature of Poly (dI-dC) ♦ Poly (dI-dC). Vibrational Circular Dichroism Results" L. Wang and T. A. Keiderling *Nucleic Acids Research* 21 4127-4132 (1993).
- "Detection and Characterization of Triple Helical Pyrimidine-Purine-Pyrimidine Nucleic Acids with Vibrational Circular Dichroism" L. Wang, P. Pancoska, T. A. Keiderling, *Biochemistry* 33 8428-8435 (1994).
- "Vibrational Circular Dichroism of A-, B- and Z- form Nucleic Acids in the PO₂- Stretching Region" L. Wang, L. Yang, T. A. Keiderling, *Biophysical Journal* 67, 2460-2467 (1994).
- "Studies of multiple stranded RNA and DNA with FTIR, vibrational and electronic circular dichroism," Zhihua Huang, Lijiang Wang and Timothy A. Keiderling, in *Spectroscopy of Biological Molecules*, Ed. J. C. Merlin, Kluwer Acad. Pub., Dordrecht, 1995, pp . 321-322.
- "Vibrational Circular Dichroism Applications to Conformational Analysis of Biomolecules" T. A. Keiderling in "Circular Dichroism and the Conformational Analysis of Biomolecules" ed G. D. Fasman, Plenum, New York (1996) pp. 555–598.
- "Vibrational Circular Dichroism Techniques and Application to Nucleic Acids" T. A. Keiderling, In "Biomolecular Structure and Dynamics", NATO ASI series, Series E: Applied Sciences- Vol.342, Eds: G. Vergoten and T. Theophanides, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 299–317 (1997).

See also

- Circular dichroism
- Birefringence
- Optical rotatory dispersion
- IR spectroscopy
- Polarization
- Proteins
- Nucleic Acids
- DNA
- Molecular models of DNA
- DNA structure
- Protein structure
- Amino acids

- Density functional theory
- Quantum chemistry
- Raman optical activity (ROA)

References

- [1] <http://planetphysics.org/?op=getobj;from=objects;id=410> Principles of IR and NIR Spectroscopy
- [2] *"Vibrational Circular Dichroism of Polypeptides XII. Re-evaluation of the Fourier Transform Vibrational Circular Dichroism of Poly-gamma-Benzyl-L-Glutamate," P. Malon, R. Kobrinskaya, T. A. Keiderling, *Biopolymers* 27, 733-746 (1988).
- [3] *"Vibrational Circular Dichroism of Biopolymers," T. A. Keiderling, S. C. Yasui, U. Narayanan, A. Annamalai, P. Malon, R. Kobrinskaya, L. Yang, in *Spectroscopy of Biological Molecules New Advances* ed. E. D. Schmid, F. W. Schneider, F. Siebert, p. 73-76 (1988).
- [4] *"Vibrational Circular Dichroism of Polypeptides and Proteins," S. C. Yasui, T. A. Keiderling, *Mikrochimica Acta*, II, 325-327, (1988).
- [5] *"Vibrational Circular Dichroism of Proteins Polysaccharides and Nucleic Acids" T. A. Keiderling, Chapter 8 in *Physical Chemistry of Food Processes, Vol. 2 Advanced Techniques, Structures and Applications.*, eds. I.C. Baianu, H. Pessen, T. Kumosinski, Van Norstrand--Reinhold, New York (1993), pp 307-337.
- [6] "Spectroscopic characterization of Unfolded peptides and proteins studied with infrared absorption and vibrational circular dichroism spectra" T. A. Keiderling and Qi Xu, *Advances in Protein Chemistry Volume 62*, [Unfolded Proteins, Dedicated to John Edsall, Ed.: George Rose, Academic Press:New York] (2002), pp. 111-161.
- [7] *"Protein and Peptide Secondary Structure and Conformational Determination with Vibrational Circular Dichroism " Timothy A. Keiderling, *Current Opinions in Chemical Biology* (Ed. Julie Leary and Mark Arnold) 6, 682-688 (2002).
- [8] *Review: Conformational Studies of Peptides with Infrared Techniques. Timothy A. Keiderling and R. A. G. D. Silva, in *Synthesis of Peptides and Peptidomimetics*, Ed. M. Goodman and G. Herrman, Houben-Weyl, Vol 22Eb, Georg Thiem Verlag, New York (2002) pp. 715-738, (written and accepted in 2000).
- [9] "Polarization Modulation Fourier Transform Infrared Spectroscopy with Digital SignalProcessing: Comparison of Vibrational Circular Dichroism Methods." Jovencio Hilario, DavidDrapcho, Raul Curbelo, Timothy A. Keiderling, *Applied Spectroscopy* 55, 1435-1447(2001)--
- [10] "Vibrational circular dichroism of biopolymers. Summary of methods and applications.", Timothy A. Keiderling, Jan Kubelka, Jovencio Hilario, in *Vibrational spectroscopy of polymers and biological systems*, Ed. Mark Braiman, Vasilis Gregoriou, Taylor&Francis, Atlanta (CRC Press, Boca Raton, FL) (2006) pp. 253-324 (originally written in 2000, updated in 2003)
- [11] "Observation of Magnetic Vibrational Circular Dichroism," T. A. Keiderling, *Journal of Chemical Physics*, 75, 3639-41 (1981).
- [12] "Vibrational Spectral Assignment and Enhanced Resolution Using Magnetic Vibrational Circular Dichroism," T. R. Devine and T. A. Keiderling, *Spectrochimica Acta*, 43A, 627-629 (1987).
- [13] "Magnetic Vibrational Circular Dichroism with an FTIR" P. V. Croatto, R. K. Yoo, T. A. Keiderling, *SPIE Proceedings* 1145 (7th International Conference on FTS, ed. D. G. Cameron) 152-153 (1989).
- [14] "Direct Measurement of the Rotational g-Value in the Ground State of Acetylene by Magnetic Vibrational Circular Dichroism." C. N. Tam and T. A. Keiderling, *Chemical Physics Letters*, 243, 55-58 (1995).
- [15] . "Ab initio calculation of the vibrational magnetic dipole moment" P. Bour, C. N. Tam, T. A. Keiderling, *Journal of Physical Chemistry* 99, 17810-17813 (1995)
- [16] "Rotationally Resolved Magnetic Vibrational Circular Dichroism. Experimental Spectra and Theoretical Simulation for Diamagnetic Molecules." P. Bour, C. N. Tam, B. Wang, T. A. Keiderling, *Molecular Physics* 87, 299-318, (1996).

Optical rotatory dispersion

Optical rotatory dispersion is the variation in the optical rotation of a substance with a change in the wavelength of light. Optical rotatory dispersion can be used to find the absolute configuration of metal complexes. For example, when plane-polarized white light from an overhead projector is passed through a cylinder of sucrose solution, a spiral rainbow is observed perpendicular to the cylinder.

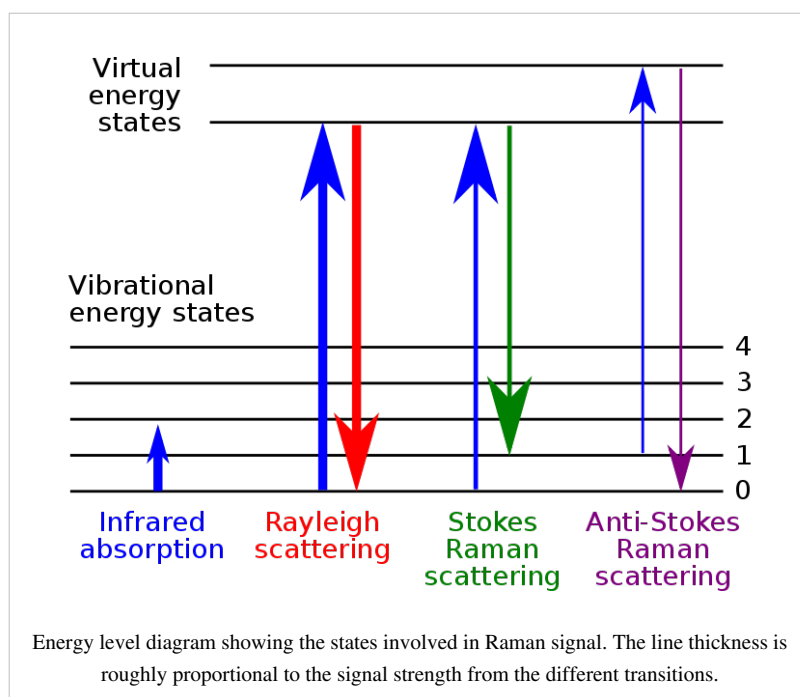
When white light passes through a polarizer, the extent of rotation of light depends on its wavelength. Short wavelengths are rotated more than longer wavelengths. Because the wavelength of light determines its color, the variation of color with distance through the tube is observed. This dependence of specific rotation on wavelength is called optical rotatory dispersion.

See also

- Circular dichroism

Raman spectroscopy

Raman spectroscopy (named after C. V. Raman, pronounced /'rɑ:mən/) is a spectroscopic technique used to study vibrational, rotational, and other low-frequency modes in a system.^[1] It relies on inelastic scattering, or Raman scattering, of monochromatic light, usually from a laser in the visible, near infrared, or near ultraviolet range. The laser light interacts with phonons or other excitations in the system, resulting in the energy of the laser photons being shifted up or down. The shift in energy gives information about the phonon modes in the system. Infrared spectroscopy yields similar, but complementary, information.



Typically, a sample is illuminated with a laser beam. Light from the illuminated spot is collected with a lens and sent through a monochromator. Wavelengths close to the laser line, due to elastic Rayleigh scattering, are filtered out while the rest of the collected light is dispersed onto a detector.

Spontaneous Raman scattering is typically very weak, and as a result the main difficulty of Raman spectroscopy is separating the weak inelastically scattered light from the intense Rayleigh scattered laser light. Historically, Raman spectrometers used holographic gratings and multiple dispersion stages to achieve a high degree of laser rejection. In the past, photomultipliers were the detectors of choice for dispersive Raman setups, which resulted in long acquisition times. However, modern instrumentation almost universally employs notch or edge filters for laser rejection and spectrographs (either axial transmissive (AT), Czerny-Turner (CT) monochromator) or FT (Fourier transform spectroscopy based), and CCD detectors.

There are a number of advanced types of Raman spectroscopy, including surface-enhanced Raman, tip-enhanced Raman, polarised Raman, stimulated Raman (analogous to stimulated emission), transmission Raman, spatially-offset Raman, and hyper Raman.

Basic theory

The Raman effect occurs when light impinges upon a molecule and interacts with the electron cloud and the bonds of that molecule. For the spontaneous Raman effect, a photon excites the molecule from the ground state to a virtual energy state. When the molecule relaxes it emits a photon and it returns to a different rotational or vibrational state. The difference in energy between the original state and this new state leads to a shift in the emitted photon's frequency away from the excitation wavelength.

If the final vibrational state of the molecule is more energetic than the initial state, then the emitted photon will be shifted to a lower frequency in order for the total energy of the system to remain balanced. This shift in frequency is designated as a Stokes shift. If the final vibrational state is less energetic than the initial state, then the emitted photon will be shifted to a higher frequency, and this is designated as an Anti-Stokes shift. Raman scattering is an example of inelastic scattering because of the energy transfer between the photons and the molecules during their interaction.

A change in the molecular polarization potential — or amount of deformation of the electron cloud — with respect to the vibrational coordinate is required for a molecule to exhibit a Raman effect. The amount of the polarizability change will determine the Raman scattering intensity. The pattern of shifted frequencies is determined by the rotational and vibrational states of the sample.

History

Although the inelastic scattering of light was predicted by Adolf Smekal in 1923, it was not until 1928 that it was observed in practice. The Raman effect was named after one of its discoverers, the Indian scientist Sir C. V. Raman who observed the effect by means of sunlight (1928, together with K. S. Krishnan and independently by Grigory Landsberg and Leonid Mandelstam).^[1] Raman won the Nobel Prize in Physics in 1930 for this discovery accomplished using sunlight, a narrow band photographic filter to create monochromatic light and a "crossed" filter to block this monochromatic light. He found that light of changed frequency passed through the "crossed" filter.

Systematic pioneering theory of the Raman effect was developed by Czechoslovak physicist George Placzek between 1930 and 1934.^[2] The mercury arc became the principal light source, first with photographic detection and then with spectrophotometric detection. Currently lasers are used as light sources.

Applications

Raman spectroscopy is commonly used in chemistry, since vibrational information is specific to the chemical bonds and symmetry of molecules. It therefore provides a fingerprint by which the molecule can be identified. For instance, the vibrational frequencies of SiO, Si₂O₂, and Si₃O₃ were identified and assigned on the basis of normal coordinate analyses using infrared and Raman spectra.^[3] The fingerprint region of organic molecules is in the (wavenumber) range 500–2000 cm⁻¹. Another way that the technique is used is to study changes in chemical bonding, e.g., when a substrate is added to an enzyme.

Raman gas analyzers have many practical applications. For instance, they are used in medicine for real-time monitoring of anaesthetic and respiratory gas mixtures during surgery.

In solid state physics, spontaneous Raman spectroscopy is used to, among other things, characterize materials, measure temperature, and find the crystallographic orientation of a sample. As with single molecules, a given solid material has characteristic phonon modes that can help an experimenter identify it. In addition, Raman spectroscopy can be used to observe other low frequency excitations of the solid, such as plasmons, magnons, and

superconducting gap excitations. The spontaneous Raman signal gives information on the population of a given phonon mode in the ratio between the Stokes (downshifted) intensity and anti-Stokes (upshifted) intensity.

Raman scattering by an anisotropic crystal gives information on the crystal orientation. The polarization of the Raman scattered light with respect to the crystal and the polarization of the laser light can be used to find the orientation of the crystal, if the crystal structure (specifically, its point group) is known.

Raman active fibers, such as aramid and carbon, have vibrational modes that show a shift in Raman frequency with applied stress. Polypropylene fibers also exhibit similar shifts. The radial breathing mode is a commonly used technique to evaluate the diameter of carbon nanotubes. In nanotechnology, a Raman microscope can be used to analyze nanowires to better understand the composition of the structures.

Spatially-offset Raman spectroscopy (SORS), which is less sensitive to surface layers than conventional Raman, can be used to discover counterfeit drugs without opening their internal packaging, and for non-invasive monitoring of biological tissue.^[4] Raman spectroscopy can be used to investigate the chemical composition of historical documents such as the Book of Kells and contribute to knowledge of the social and economic conditions at the time the documents were produced.^[5] This is especially helpful because Raman spectroscopy offers a non-invasive way to determine the best course of preservation or conservation treatment for such materials.

Raman spectroscopy is being investigated as a means to detect explosives for airport security.^[6]

Microspectroscopy

Raman spectroscopy offers several advantages for microscopic analysis. Since it is a scattering technique, specimens do not need to be fixed or sectioned. Raman spectra can be collected from a very small volume ($< 1 \mu\text{m}$ in diameter); these spectra allow the identification of species present in that volume. Water does not generally interfere with Raman spectral analysis. Thus, Raman spectroscopy is suitable for the microscopic examination of minerals, materials such as polymers and ceramics, cells and proteins. A Raman microscope begins with a standard optical microscope, and adds an excitation laser, a monochromator, and a sensitive detector (such as a charge-coupled device (CCD), or photomultiplier tube (PMT)). FT-Raman has also been used with microscopes.

In *direct imaging*, the whole field of view is examined for scattering over a small range of wavenumbers (Raman shifts). For instance, a wavenumber characteristic for cholesterol could be used to record the distribution of cholesterol within a cell culture.

The other approach is *hyperspectral imaging* or *chemical imaging*, in which thousands of Raman spectra are acquired from all over the field of view. The data can then be used to generate images showing the location and amount of different components. Taking the cell culture example, a hyperspectral image could show the distribution of cholesterol, as well as proteins, nucleic acids, and fatty acids. Sophisticated signal- and image-processing techniques can be used to ignore the presence of water, culture media, buffers, and other interferents.

Raman microscopy, and in particular confocal microscopy, has very high spatial resolution. For example, the lateral and depth resolutions were 250 nm and 1.7 μm , respectively, using a confocal Raman microspectrometer with the 632.8 nm line from a He-Ne laser with a pinhole of 100 μm diameter. Since the objective lenses of microscopes focus the laser beam to several micrometres in diameter, the resulting photon flux is much higher than achieved in conventional Raman setups. This has the added benefit of enhanced fluorescence quenching. However, the high photon flux can also cause sample degradation, and for this reason some setups require a thermally conducting substrate (which acts as a heat sink) in order to mitigate this process.

By using Raman microspectroscopy, *in vivo* time- and space-resolved Raman spectra of microscopic regions of samples can be measured. As a result, the fluorescence of water, media, and buffers can be removed. Consequently *in vivo* time- and space-resolved Raman spectroscopy is suitable to examine proteins, cells and organs.

Raman microscopy for biological and medical specimens generally uses near-infrared (NIR) lasers (785 nm diodes and 1064 nm Nd:YAG are especially common). This reduces the risk of damaging the specimen by applying higher

energy wavelengths. However, the intensity of NIR Raman is low (owing to the ω^4 dependence of Raman scattering intensity), and most detectors required very long collection times. Recently, more sensitive detectors have become available, making the technique better suited to general use. Raman microscopy of inorganic specimens, such as rocks and ceramics and polymers, can use a broader range of excitation wavelengths.^[7]

Polarized analysis

The polarization of the Raman scattered light also contains useful information. This property can be measured using (plane) polarized laser excitation and a polarization analyzer. Spectra acquired with the analyzer set at both perpendicular and parallel to the excitation plane can be used to calculate the depolarization ratio. Study of the technique is pedagogically useful in teaching the connections between group theory, symmetry, Raman activity and peaks in the corresponding Raman spectra.

The spectral information arising from this analysis gives insight into molecular orientation and vibrational symmetry. In essence, it allows the user to obtain valuable information relating to the molecular shape, for example in synthetic chemistry or polymorph analysis. It is often used to understand macromolecular orientation in crystal lattices, liquid crystals or polymer samples.^[8]

Variations

Several variations of Raman spectroscopy have been developed. The usual purpose is to enhance the sensitivity (e.g., surface-enhanced Raman), to improve the spatial resolution (Raman microscopy), or to acquire very specific information (resonance Raman).

- **Surface Enhanced Raman Spectroscopy (SERS)** - Normally done in a silver or gold colloid or a substrate containing silver or gold. Surface plasmons of silver and gold are excited by the laser, resulting in an increase in the electric fields surrounding the metal. Given that Raman intensities are proportional to the electric field, there is large increase in the measured signal (by up to 10^{11}). This effect was originally observed by Martin Fleischmann but the prevailing explanation was proposed by Van Duyne in 1977.^[9]
- **Resonance Raman spectroscopy** - The excitation wavelength is matched to an electronic transition of the molecule or crystal, so that vibrational modes associated with the excited electronic state are greatly enhanced. This is useful for studying large molecules such as polypeptides, which might show hundreds of bands in "conventional" Raman spectra. It is also useful for associating normal modes with their observed frequency shifts.^[10]
- **Surface Enhanced Resonance Raman Spectroscopy (SERRS)** - A combination of SERS and resonance Raman spectroscopy which uses proximity to a surface to increase Raman intensity, and excitation wavelength matched to the maximum absorbance of the molecule being analysed.
- **Hyper Raman** - A non-linear effect in which the vibrational modes interact with the second harmonic of the excitation beam. This requires very high power, but allows the observation of vibrational modes which are normally "silent". It frequently relies on SERS-type enhancement to boost the sensitivity.^[11]
- **Spontaneous Raman Spectroscopy** - Used to study the temperature dependence of the Raman spectra of molecules.
- **Optical Tweezers Raman Spectroscopy (OTRS)** - Used to study individual particles, and even biochemical processes in single cells trapped by optical tweezers.
- **Stimulated Raman Spectroscopy** - A spatially coincident, two color pulse (with polarization either parallel or perpendicular) transfers the population from ground to a rovibrationally excited state, if the difference in energy corresponds to an allowed Raman transition, and if neither frequency corresponds to an electronic resonance. Two photon UV ionization, applied after the population transfer but before relaxation, allows the intra-molecular or inter-molecular Raman spectrum of a gas or molecular cluster (indeed, a given conformation of molecular cluster) to be collected. This is a useful molecular dynamics technique.

- **Spatially Offset Raman Spectroscopy (SORS)** - The Raman scatter is collected from regions laterally offset away from the excitation laser spot, leading to significantly lower contributions from the surface layer than with traditional Raman spectroscopy.^[12]
- **Coherent anti-Stokes Raman spectroscopy (CARS)** - Two laser beams are used to generate a coherent anti-Stokes frequency beam, which can be enhanced by resonance.
- **Raman optical activity (ROA)** - Measures vibrational optical activity by means of a small difference in the intensity of Raman scattering from chiral molecules in right- and left-circularly polarized incident light or, equivalently, a small circularly polarized component in the scattered light.^[13]
- **Transmission Raman** - Allows probing of a significant bulk of a turbid material, such as powders, capsules, living tissue, etc. It was largely ignored following investigations in the late 1960s^[14] but was rediscovered in 2006 as a means of rapid assay of pharmaceutical dosage forms.^[15] There are also medical diagnostic applications.^[16]
- **Inverse Raman spectroscopy.**
- **Tip-Enhanced Raman Spectroscopy (TERS)** - Uses a metallic (usually silver-/gold-coated AFM or STM) tip to enhance the Raman signals of molecules situated in its vicinity. The spatial resolution is approximately the size of the tip apex (20-30 nm). TERS has been shown to have sensitivity down to the single molecule level.

External links

- An introduction to Raman spectroscopy^[17], horiba.com
- Raman Application examples^[18], horiba.com
- A introduction on Raman Scattering^[19], d3technologies.co.uk
- Raman Spectroscopy Applications^[20], renishaw.com
- Romanian Database of Raman Spectroscopy^[21] - This database contains mineral species (natural and synthetic) with description of crystal structure, sample image, number of sample, origin, Raman spectrum and vibrations, Raman discussion and references. Also, this site contains artefacts sample with sample image and pigment spectrum; black, red, white or blue pigment. rdrs.uaic.ro
- Chemical Imaging Without Dyeing^[22], witec.de
- DoITPoMS Teaching and Learning Package - Raman Spectroscopy^[23] - an introduction, aimed at undergraduate level. doitpoms.ac.uk
- Raman Spectroscopy Tutorial^[24] - A detailed explanation of Raman Spectroscopy including Resonance-Enhanced Raman Scattering and Surface-Enhanced Raman Scattering. 161.25.205.25
- The Science Show, ABC Radio National^[25] - Interview with Scientist on NASA funded project to build Raman Spectrometer for the 2009 Mars mission: a cellular phone size device to detect almost any substance known, with commercial <USD\$5000 commercial spin-off, prototyped by June 2006. abc.net.au/rn
- Raman spectroscopy for medical diagnosis^[26] from the June 1, 2007 issue of *Analytical Chemistry*^[27], pubs.acs.org
- Spontaneous Raman Scattering (SRS)^[28], lavision.de

References

- [1] Gardiner, D.J. (1989). *Practical Raman spectroscopy*. Springer-Verlag. ISBN 978-0387502540.
- [2] Placzek G.: "Rayleigh Streueung und Raman Effekt", In: Hdb. der Radiologie, Vol. VI., 2, 1934, p. 209
- [3] Khanna, R.K. (1981). "Raman-spectroscopy of oligomeric SiO species isolated in solid methane". *Journal of Chemical Physics*. doi:10.1063/1.441393.
- [4] . BBC News. 2007-01-31. <http://news.bbc.co.uk/2/hi/health/6314287.stm>. Retrieved 2008-12-08.
- [5] Irish classic is still a hit (in calfskin, not paperback) - New York Times (<http://www.nytimes.com/2007/05/28/world/europe/28kells.html>), nytimes.com
- [6] Ben Vogel (29 August 2008). "Raman spectroscopy portends well for standoff explosives detection" (http://www.janes.com/news/transport/business/jar/jar080829_1_n.shtml). Jane's. . Retrieved 2008-08-29.
- [7] Ellis DI, Goodacre R (August 2006). "Metabolic fingerprinting in disease diagnosis: biomedical applications of infrared and Raman spectroscopy". *Analyst* **131** (8): 875–85. doi:10.1039/b602376m. PMID 17028718.
- [8] Khanna, R.K. (1957). *Evidence of ion-pairing in the polarized Raman spectra of a Ba2+CrO doped KI single crystal*. John Wiley & Sons, Ltd. doi:10.1002/jrs.1250040104.
- [9] Jeanmaire DL, van Duyne RP (1977). "Surface Raman Electrochemistry Part I. Heterocyclic, Aromatic and Aliphatic Amines Adsorbed on the Anodized Silver Electrode". *Journal of Electroanalytical Chemistry* (Elsevier Sequouia S.A.) **84**: 1–20. doi:10.1016/S0022-0728(77)80224-6.
- [10] Chao RS, Khanna RK, Lippincott ER (1974). "Theoretical and experimental resonance Raman intensities for the manganate ion". *J Raman Spectroscopy* **3**: 121. doi:10.1002/jrs.1250030203.
- [11] Kneipp K, *et al.* (1999). "Surface-Enhanced Non-Linear Raman Scattering at the Single Molecule Level". *Chem. Phys.* **247**: 155–162. doi:10.1016/S0301-0104(99)00165-2.
- [12] Matousek P, Clark IP, Draper ERC, *et al.* (2005). "Subsurface Probing in Diffusely Scattering Media using Spatially Offset Raman Spectroscopy". *Applied Spectroscopy* **59**: 393. doi:10.1366/000370205775142548.
- [13] Barron LD, Hecht L, McColl IH, Blanch EW (2004). "Raman optical activity comes of age". *Molec. Phys.* **102** (8): 731–744. doi:10.1080/00268970410001704399.
- [14] B. Schrader, G. Bergmann, Fresenius. Z. (1967). *Anal. Chem.*: 225–230.
- [15] P. Matousek, A. W. Parker (2006). "Bulk Raman Analysis of Pharmaceutical Tablets". *Applied Spectroscopy* **60**: 1353–1357. doi:10.1366/000370206779321463.
- [16] P. Matousek, N. Stone (2007). "Prospects for the diagnosis of breast cancer by noninvasive probing of calcifications using transmission Raman spectroscopy". *Journal of Biomedical Optics* **12**: 024008. doi:10.1117/1.2718934.
- [17] <http://www.horiba.com/us/en/scientific/products/raman-spectroscopy/raman-resource/raman-tutorial/>
- [18] <http://www.horiba.com/scientific/products/raman-spectroscopy/application-notes/>
- [19] <http://www.d3technologies.co.uk/en/10371.aspx>
- [20] <http://www.renishaw.com/en/raman-spectroscopy-applications--6259>
- [21] <http://rdrs.uaic.ro>
- [22] <http://witec.de/en/download/Raman/ImagingMicroscopy04.pdf>
- [23] <http://www.doitpoms.ac.uk/tlplib/raman/index.php>
- [24] http://161.58.205.25/Raman_Spectroscopy/rtr-ramantutorial.php?ss=800
- [25] <http://www.abc.net.au/rn/science/ss/stories/s1581469.htm>
- [26] http://pubs.acs.org/subscribe/journals/ancham/79/i11/pdf/0607feature_griffiths.pdf
- [27] <http://pubs3.acs.org/acs/journals/toc.page?incoden=ancham&indecade=0&involume=79&inissue=11>
- [28] http://www.lavision.de/techniques/mie_raman_rayleigh/raman_imaging.php

Coherent anti-Stokes Raman spectroscopy

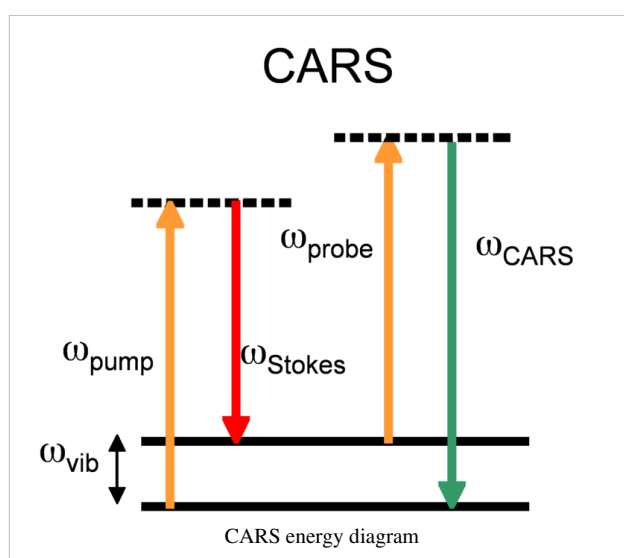
Coherent anti-Stokes Raman spectroscopy, also called Coherent anti-Stokes Raman scattering spectroscopy (CARS), is a form of spectroscopy used primarily in chemistry, physics and related fields. It is sensitive to the same vibrational signatures of molecules as seen in Raman spectroscopy, typically the nuclear vibrations of chemical bonds. Unlike Raman spectroscopy, CARS employs multiple photons to address the molecular vibrations, and produces a signal in which the emitted waves are coherent with one another. As a result, CARS is orders of magnitude stronger than spontaneous Raman emission. CARS is a third-order nonlinear optical process involving three laser beams: a pump beam of frequency ω_p , a Stokes beam of frequency ω_s and a probe beam at frequency ω_{pr} . These beams interact with the sample and generate a coherent optical signal at the anti-Stokes frequency ($\omega_p - \omega_s + \omega_{pr}$). The latter is resonantly enhanced when the frequency difference between the pump and the Stokes beams ($\omega_p - \omega_s$) coincides with the frequency of a Raman resonance, which is the basis of the technique's intrinsic vibrational contrast mechanism.^{[1] [2]}

History

The acronym CARS, which invokes a seemingly inadvertent relation to automobiles, is actually closely related to the birth story of the technique. In 1965, a paper was published by two researchers of the Scientific Laboratory at the Ford Motor Company, P. D. Maker and R. W. Terhune, in which the CARS phenomenon was reported for the first time.^[3] Maker and Terhune used a pulsed ruby laser to investigate the third order response of several materials. They first passed the ruby beam of frequency ω through a Raman shifter to create a second beam at $\omega - \omega_v$, and then directed the two beams simultaneously onto the sample. When the pulses from both beams overlapped in space and time, the Ford researchers observed a signal at $\omega + \omega_v$, which is the blue-shifted CARS signal. They also demonstrated that the signal increases significantly when the difference frequency ω_v between the incident beams matches a Raman frequency of sample. Maker and Terhune called their technique simply 'three wave mixing experiments'. The name coherent anti-Stokes Raman spectroscopy was assigned almost ten years later, by Begley et al. at Stanford University in 1974.^[4] Since then, this vibrationally sensitive nonlinear optical technique is commonly known as CARS.

Principle

The CARS process can be physically explained by using either a classical oscillator model or by using a quantum mechanical model that incorporates the energy levels of the molecule. Classically, the Raman active vibrator is modeled as a (damped) harmonic oscillator with a characteristic frequency of ω_v . In CARS, this oscillator is not driven by a single optical wave, but by the difference frequency ($\omega_p - \omega_s$) between the pump and the Stokes beams instead. This driving mechanism is similar to hearing the low combination tone when striking two different high tone piano keys: your ear is sensitive to the difference frequency of the high tones. Similarly, the Raman oscillator is susceptible to the difference frequency of two optical



waves. When the difference frequency $\omega_p - \omega_s$ approaches ω_v , the oscillator is driven very efficiently. On a molecular level, this implies that the electron cloud surrounding the chemical bond is vigorously oscillating with the frequency $\omega_p - \omega_s$. These electron motions alter the optical properties of the sample, i.e. there is a periodic modulation of the refractive index of the material. This periodic modulation can be probed by a third laser beam, the probe beam. When the probe beam is propagating through the periodically altered medium, it acquires the same modulation. Part of the probe, originally at ω_{pr} will now get modified to $\omega_{pr} + \omega_p - \omega_s$, which is the observed anti-Stokes emission. Under certain beam geometries, the anti-Stokes emission may diffract away from the probe beam, and can be detected in a separate direction.

While intuitive, this classical picture does not take into account the quantum mechanical energy levels of the molecule. Quantum mechanically, the CARS process can be understood as follows. Our molecule is initially in the ground state, the lowest energy state of the molecule. The pump beam excites the molecule to a Virtual State. A virtual state is not an eigenstate of the molecule and it can not be occupied but it does allow for transitions between otherwise uncoupled real states. If a Stokes beam is simultaneously present along with the pump, the virtual state can be used as an instantaneous gateway to address a vibrational eigenstate of the molecule. The joint action of the pump and the Stokes has effectively established a coupling between the ground state and the vibrationally excited state of the molecule. The molecule is now in two states at the same time: it resides in a coherent superposition of states. This coherence between the states can be probed by the probe beam, which promotes the system to a virtual state. Again, the molecule cannot stay in the virtual state and will fall back instantaneously to the ground state under the emission of a photon at the anti-Stokes frequency. The molecule is no longer in a superposition, as it resides again in one state, the ground state. In the quantum mechanical model, no energy is deposited in the molecule during the CARS process. Instead, the molecule acts like a medium for converting the frequencies of the three incoming waves into a CARS signal (a parametric process). There are, however, related coherent Raman process that occur simultaneously which do deposit energy into the molecule.

Comparison to Raman spectroscopy

CARS is often compared to Raman spectroscopy as both techniques probe the same Raman active modes. Raman can be done using a single continuous wave (CW) laser whereas CARS requires (generally) two pulsed laser sources. The Raman signal is detected on the red side of the incoming radiation where it might have to compete with other fluorescent processes. The CARS signal is detected on the blue side, which is free from fluorescence, but it comes with a non-resonant contribution. The differences between the signals from Raman and CARS (there are many variants of both techniques) stems largely from the fact that Raman relies on a spontaneous transition whereas CARS relies on a coherently driven transition. The total Raman signal collected from a sample is the incoherent addition of the signal from individual molecules. It is therefore linear in the concentration of those molecules and the signal is emitted in all directions. The total CARS signal comes from a coherent addition of the signal from individual molecules. For the coherent addition to be additive, phase-matching must be fulfilled. For tight focusing conditions this is generally not a restriction. Once phase-matching is fulfilled the signal amplitude grows linear with distance so that the power grows quadratically. This signal forms a collimated beam that is therefore easily collected. The fact that the CARS signal is quadratic in the distance makes it quadratic with respect to the concentration and therefore especially sensitive to the majority constituent. The total CARS signal also contains an inherent non-resonant background. This non-resonant signal can be considered as the result of (several) far off-resonance transitions that also add coherently. The resonant amplitude contains a phase shift of π over the resonance whereas the non-resonant part does not. The spectral line shape of the CARS intensity therefore resembles a Fano-profile which is shifted with respect to the Raman signal. To compare the spectra from multi-component compounds, the (resonant) CARS spectral amplitude should be compared to the Raman spectral intensity.

Theoretically Raman spectroscopy and CARS spectroscopy are equally sensitive as they use the same molecular transitions. However, given the limits on input power (damage threshold) and detector noise (integration time), the

signal from a single transition can be collected much faster in practical situation (a factor of 10^5) using CARS. Imaging of known substances (known spectra) is therefore often done using CARS. Given the fact that CARS is a higher order nonlinear process, the CARS signal from a single molecule is larger than the Raman signal from a single molecule for a sufficiently high driving intensity. However at very low concentrations, the advantages of the coherent addition for CARS signal reduces and the presence of the incoherent background becomes an increasing problem.

Since CARS is such a nonlinear process there are not really any 'typical' experimental numbers. One example is given below under the explicit warning that just changing the pulse duration by one order of magnitude changes the CARS signal by three orders of magnitude. The comparison should only be used as an indication of the order of magnitude of the signals. 200 mW average power input (CW for the Raman), in a 0.9NA objective with a center wavelength around 800 nm, constitutes a power density of 26 MW/cm^2 (focal length = 1.5 micrometre, focal volume = $1.16 \text{ micrometre}^3$, photon energy = $2.31 \cdot 10^{-19} \text{ J}$ or 1.44 eV). The Raman cross section for the vibration of the aromatic ring in Toluene around 1000 cm^{-1} is on the order of $10^{-29} \text{ cm}^2/\text{molecule} \cdot \text{steradian}$. Therefore the Raman signal is around $26 \cdot 10^{-22} \text{ W/molecule} \cdot \text{steradian}$ or $3.3 \cdot 10^{-21} \text{ W/molecule}$ (over 4π). That is $0.014 \text{ photon/sec} \cdot \text{molecule}$. The density of Toluene = $0.8668 \cdot 10^3 \text{ kg/m}^3$, Molecular mass = $92.14 \cdot 10^{-3} \text{ kg/mol}$. Therefore the focal volume ($\sim 1 \text{ cubic micrometre}$) contains $6 \cdot 10^9$ molecules. Those molecules together generate a Raman signal in the order of $2 \cdot 10^{-11} \text{ W}$ (20pW) or roughly one hundred million photons/sec (over a 4π solid angle). A CARS experiment with similar parameters (150 mW at 1064 nm, 200 mW at 803.5 nm, 15ps pulses at 80MHz repetition frequency, same objective lens) yields roughly $17.5 \cdot 10^{-6} \text{ W}$ (on the 3000 cm^{-1} line, which has 1/3 of the strength and roughly 3 times the width). This CARS power is roughly 10^6 higher than the Raman but since there are $6 \cdot 10^9$ molecules, the signal per molecule from CARS is only $4 \cdot 10^{-25} \text{ W/molecule} \cdot \text{sec}$ or $1.7 \cdot 10^{-6} \text{ photons/molecule} \cdot \text{sec}$. If we allow for two factors of three (line strength and line width) than the spontaneous Raman signal per molecule still exceeds the CARS per molecule by a more than two orders of magnitude. The coherent addition of the CARS signal from the molecules however yields a total signal that is much higher than the Raman.

The sensitivity in many CARS experiments is not limited by the detection of CARS photons but rather by the distinction between the resonant and non-resonant part of the CARS signal.

Applications

CARS is used for species selective microscopy and combustion diagnostics. The first exploits the selectivity of vibrational spectroscopy whereas the latter is aimed at temperature measurements; the CARS signal is temperature dependent. The strength of the signal scales (non-linearly) with the difference in the ground state population and the vibrationally excited state population. Since the population of states follows the temperature dependent Boltzmann Distribution, the CARS signal carries an intrinsic temperature dependence as well. This temperature dependence makes CARS a popular technique for monitoring the temperature of hot gases and flames.

See also

- Coherent Stokes Raman spectroscopy
- Raman spectroscopy
- Four-wave mixing

References

- [1] A Review of the Theory and Application of Coherent Anti-Stokes Raman Spectroscopy (CARS) *Applied Spectroscopy*, Volume 31, Number 4, July/August 1977, pp. 253-271(19) (<http://www.ingentaconnect.com/content/sas/sas/1977/00000031/00000004/art00001>)
- [2] Coherent anti-Stokes Raman scattering: from proof-of-the-principle experiments to femtosecond CARS and higher order wave-mixing generalizations *Journal of Raman Spectroscopy*, Volume 31, Issue 8-9, pp. 653 - 667 (<http://www3.interscience.wiley.com/cgi-bin/abstract/73500427/ABSTRACT?CRETRY=1&SRETRY=0>)
- [3] Study of Optical Effects Due to an Induced Polarization Third Order in the Electric Field Strength *Physical Review*, Volume 137, Issue 3A, pp. 801-818 (http://prola.aps.org/abstract/PR/v137/i3A/pA801_1)
- [4] Coherent anti-Stokes Raman spectroscopy *Applied Physics Letters*, Volume 25, Issue 7, pp. 387-390 (<http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=APPLAB000025000007000387000001&idtype=cvips&gifs=yes>)

Raman Microscopy

Raman microscope begins with a standard optical microscope, and adds an excitation laser, a monochromator, and an optical sensitive detector such as a charge-coupled device (CCD), or photomultiplier tube, (PMT)), and has been implemented for Raman spectroscopy in direct chemical imaging, the whole field of view on 3D sample.

Imaging spectroscopy

Imaging spectroscopy (also **spectral imaging**, **chemical imaging**, or **microspectroscopy**) is similar to color photography, but each pixel acquires many bands of light intensity data from the spectrum, instead of just the three bands of the RGB color model. More precisely, it is the simultaneous acquisition of spatially coregistered images in many spectrally contiguous bands.



Ash plumes on Kamchatka Peninsula, eastern Russia. A MODIS image.

Some spectral images contain only a few image planes of spectral data, while others are better thought of as full spectra at every location in the image. For example, solar physicists use the spectroheliograph, to make images of the Sun built up by scanning the slit of a spectrograph, to study the behavior of surface features on the Sun; such a spectroheliogram may have a spectral resolution of over 100,000 ($\lambda/\Delta\lambda$) and be used to measure local motion (via the Doppler shift) and even the magnetic field (via the Zeeman splitting or Hanle effect) at each location in the image plane. The multispectral images collected by the Opportunity rover, in contrast, have only four wavelength

bands and hence are only a little more than 3-color images.

To be scientifically useful, such measurement should be done using an internationally recognized system of units.

One example application is geophysical spectral imaging, which allows quantitative and qualitative characterization of both, the surface and the atmosphere, using geometrically coherent spectrodirectional radiometric measurements. These measurements can then be used for the unambiguous direct and indirect identification of surface materials and atmospheric trace gases, the measurement of their relative concentrations, subsequently the assignment of the proportional contribution of mixed pixel signals (e.g., the spectral unmixing problem), the derivation of their spatial distribution (mapping problem), and finally their study over time (multi-temporal analysis). The Moon Mineralogy Mapper on Chandrayaan-1 was an imaging spectrometer.^[1]

Background

About 300 years ago, in 1704, Sir Isaac Newton published in his 'Treatise of Light' (Newton, 1704) the concept of dispersion of light. He demonstrated that white light could be split up into component colours by means of a prism, and found that each pure colour is characterized by a specific refrangibility. The corpuscular theory by Newton was gradually succeeded over time by the wave theory. Consequently, the substantial summary of past experiences performed by Maxwell (1873), resulted in his equations of electromagnetic waves. But it was not until the 19th century that the quantitative measurement of dispersed light was recognized and standardized.

A major contribution was Fraunhofer's discovery of the dark lines in the solar spectrum (Fraunhofer, 1817); and their interpretation as absorption lines on the basis of experiments by Bunsen and Kirchhoff (1863). The term "spectroscopy" was first used in the late 19th century and provides the empirical foundations for atomic and molecular physics (Born & Wolf, 1999). Significant achievements in imaging spectroscopy are attributed to airborne instruments, particularly arising in the early 1980s and 1990s (Goetz et al., 1985; Vane et al., 1984). However, it was not until 1999 that the first imaging spectrometer was launched in space (the NASA Moderate-resolution Imaging Spectroradiometer, or MODIS).

Terminology and definitions evolve over time. At one time, >10 spectral bands sufficed to justify the term "imaging spectrometer" but presently the term is seldom defined by a total minimum number of spectral bands, rather by a contiguous (or redundant) statement of spectral bands.

The term hyperspectral imaging is sometimes used interchangeably with imaging spectroscopy. Due to its heavy use in military related applications, the civil world has established a slight preference for using the term imaging spectroscopy.

Unmixing

Hyperspectral data is often used to determine what materials are present in a scene. Materials of interest could include roadways, vegetation, and specific targets (i.e. pollutants, hazardous materials, etc). Trivially, each pixel of a hyperspectral image could be compared to a material database to determine the type of material making up the pixel. However, many hyperspectral imaging platforms have low resolution (>5m per pixel) causing each pixel to be a mixture of several materials. The process of unmixing one of these 'mixed' pixels is called hyperspectral image unmixing or simply hyperspectral unmixing.

Models

A solution to hyperspectral unmixing is to reverse the mixing process. Generally, two models of mixing are assumed: linear and nonlinear. Linear mixing models the ground as being flat and incident sunlight on the ground causes the materials to radiate some amount of the incident energy back to the sensor. Each pixel then, is modeled as a linear sum of all the radiated energy curves of materials making up the pixel. Therefore, each material contributes to the sensor's observation in a positive linear fashion. Additionally, a conservation of energy constraint is often

observed thereby forcing the weights of the linear mixture to sum to one in addition to being positive. The model can be described mathematically as follows:

$$p = A * x$$

where p represents a pixel observed by the sensor, A is a matrix of material reflectance signatures (each signature is a column of the matrix), and x is the proportion of material present in the observed pixel. This type of model is also referred to as a simplex.

With x satisfying the two constraints: 1. Abundance Nonnegativity Constraint (ANC) - each element of x is positive. 2. Abundance Sum-to-one Constraint (ASC) - the elements of x must sum to one.

Non-linear mixing results from multiple scattering often due to non-flat surface such as buildings and vegetation.

Unmixing (Endmember Detection) Algorithms

There are many algorithms to unmix hyperspectral data each with their own strengths and weaknesses. Many algorithms assume that pure pixels (pixels which contain only one materials) are present in a scene. Some algorithms to perform unmixing are listed below:

- Pixel Purity Index (PPI) - Works by projecting each pixel onto one vector from a set of random vectors spanning the reflectance space. A pixel receives a score when it represent an extremum of all the projections. Pixels with the highest scores are deemed to be spectrally pure.
- NFINDR
- Gift Wrapping Algorithm
- Independent Component Analysis Endmember Extraction Algorithm (ICA-EEA) - Works by assuming that pure pixels occur independently than mixed pixels. Assumes pure pixels are present.
- Vertex Component Analysis (VCA) - Works on the fact that the affine transformation of a simplex is another simplex which helps to find hidden (folded) vertices of the simplex. Assumes pure pixels are present.
- Principal component analysis -(PCA) could also be used to determine endmembers, projection on principal axes could permit endmember selection [Smith,Johnson et Adams (1985), Bateson et Curtiss (1996)]
- Multi Endmembers Spatial Mixture Analysis (MESMA) based on the SMA algorithm

Non-linear unmixing algorithm also exist (Support Vector Machines (SVM)) or Analytical Neural Network (ANN). Probabilistics methods have also been attempted to unmix pixel through Monte Carlo Unmixing (MCU) algorithm.

Abundance Maps

Once the fundamental materials of a scene are determined, it is often useful to construct an abundance map of each material which displays the fractional amount of material present at each pixel. Often linear programming is done to observed ANC and ASC.

Sensors

- AVIRIS — aircraft, (224 bands)
- Telops Hyper-Cam^[2] — Commercial infrared hyperspectral camera, ground-based or aircraft
- MODIS — on board EOS Terra and Aqua platforms
- MERIS — on board Envisat
- HyDice
- Hyperion — on board the Earth Observing 1 satellite

See also

- Remote sensing
- Hyperspectral
- Full Spectral Imaging
- List of Earth observation satellites
- Chemical Imaging
- Imaging spectrometer
- Infrared Microscopy

References

- [1] Large quantities of water found on the Moon (<http://www.telegraph.co.uk/science/space/6224974/Large-quantities-of-water-found-on-the-Moon.html>)
- [2] <http://www.hyper-cam.com/>
- Born, M. & Wolf, E. (1999) *Principles of Optics*, 7 edn. Cambridge University Press, Cambridge.
 - Bunsen, R. & Kirchhoff, G. (1863) *Untersuchungen über das Sonnenspektrum und die Spektren der Chemischen Elemente*. Abh. kgl. Akad. Wiss., 1861.
 - Fraunhofer, J. (1817) Bestimmung des Brechungs- und Farbenzerstreuungs-Vermögens verschiedener Glasarten, in Bezug auf die Vervollkommnung achromatischer Fernrohre, Vol. 56, pp. 264–313. *Gilberts Annalen der Physik*.
 - Goetz, A.F.H., Vane, G., Solomon, J.E., & Rock, B.N. (1985) Imaging spectrometry for earth remote sensing. *Science*, 228, 1147.
 - Maxwell, J.C. (1873) *A Treatise on Electricity and Magnetism* Clarendon Press, Oxford.
 - Newton, I. (1704) *Opticks: Or, a Treatise of the Reflexions, Refractions, Inflexions and Colours of Light* Sam Smith and Benj. Walford, London.
 - Schaepman, M. (2005) *Spectrodirectional Imaging: From Pixels to Processes*. Inaugural address, Wageningen University, Wageningen (NL).
 - Vane, G., Chrisp, M., Emmark, H., Macenka, S., & Solomon, J. (1984) *Airborne Visible Infrared Imaging Spec-trometer (AVIRIS): An Advanced Tool for Earth Remote Sensing*. European Space Agency, (Special Publication) ESA SP, 2, 751.

External links

- About imaging spectroscopy (USGS): <http://speclab.cr.usgs.gov/aboutimsp.html>
 - Link to resources (OKSI): http://www.techexpo.com/WWW/opto-knowledge/IS_resources.html
 - Special Interest Group Imaging Spectroscopy (EARSel): <http://www.op.dlr.de/dais/SIG-IS/SIG-IS.html>
 - Applications of Spectroscopic and Chemical Imaging in Research: <http://www3.imperial.ac.uk/vibrationalspectroscopyandchemicalimaging/research>
-

Chemical imaging

Chemical imaging is the analytical capability (as quantitative - mapping) to create a visual image from simultaneous measurement of spectra (as quantitative - chemical) and spatial, time informations.^{[1] [2]} The technique is most often applied to either solid or gel samples, and has applications in chemistry, biology^{[3] [4] [5] [6] [7] [8]}, medicine^{[9] [10]}, pharmacy^[11] (see also for example: Chemical Imaging Without Dyeing^[22]), food science, biotechnology^{[12] [13]}, agriculture and industry (see for example: NIR Chemical Imaging in Pharmaceutical Industry^[14] and Pharmaceutical Process Analytical Technology: ^[15]). NIR, IR and Raman chemical imaging is also referred to as hyperspectral, spectroscopic, spectral or multispectral imaging (also see microspectroscopy). However, other ultra-sensitive and selective, chemical imaging techniques are also in use that involve either UV-visible or fluorescence microspectroscopy. Chemical imaging techniques can be used to analyze samples of all sizes, from the single molecule^{[16] [17]} to the cellular level in biology and medicine^{[18] [19] [20]}, and to images of planetary systems in astronomy, but different instrumentation is employed for making observations on such widely different systems.

Chemical imaging instrumentation is composed of three components: a radiation source to illuminate the sample, a spectrally selective element, and usually a detector array (the camera) to collect the images. When many stacked spectral channels (wavelengths) are collected for different locations of the microspectrometer focus on a line or planar array in the focal plane, the data is called hyperspectral; fewer wavelength data sets are called multispectral. The data format is called a hypercube. The data set may be visualized as a three-dimensional block of data spanning two spatial dimensions (x and y), with a series of wavelengths (lambda) making up the third (spectral) axis. The hypercube can be visually and mathematically treated as a series of spectrally resolved images (each image plane corresponding to the image at one wavelength) or a series of spatially resolved spectra. The analyst may choose to view the spectrum measured at a particular spatial location; this is useful for chemical identification. Alternatively, selecting an image plane at a particular wavelength can highlight the spatial distribution of sample components, provided that their spectral signatures are different at the selected wavelength.

Many materials, both manufactured and naturally occurring, derive their functionality from the spatial distribution of sample components. For example, extended release pharmaceutical formulations can be achieved by using a coating that acts as a barrier layer. The release of active ingredient is controlled by the presence of this barrier, and imperfections in the coating, such as discontinuities, may result in altered performance. In the semi-conductor industry, irregularities or contaminants in silicon wafers or printed micro-circuits can lead to failure of these components. The functionality of biological systems is also dependent upon chemical gradients – a single cell, tissue, and even whole organs function because of the very specific arrangement of components. It has been shown that even small changes in chemical composition and distribution may be an early indicator of disease.

Any material that depends on chemical gradients for functionality may be amenable to study by an analytical technique that couples spatial and chemical characterization. To efficiently and effectively design and manufacture such materials, the 'what' and the 'where' must both be measured. The demand for this type of analysis is increasing as manufactured materials become more complex. Chemical imaging techniques not only permit visualization of the spatially resolved chemical information that is critical to understanding modern manufactured products, but it is also a non-destructive technique so that samples are preserved for further testing.

History

Commercially available laboratory-based chemical imaging systems emerged in the early 1990s (ref. 1-5). In addition to economic factors, such as the need for sophisticated electronics and extremely high-end computers, a significant barrier to commercialization of infrared imaging was that the focal plane array (FPA) needed to read IR images were not readily available as commercial items. As high-speed electronics and sophisticated computers became more commonplace, and infrared cameras became readily commercially available, laboratory chemical imaging systems were introduced.

Initially used for novel research in specialized laboratories, chemical imaging became a more commonplace analytical technique used for general R&D, quality assurance (QA) and quality control (QC) in less than a decade. The rapid acceptance of the technology in a variety of industries (pharmaceutical, polymers, semiconductors, security, forensics and agriculture) rests in the wealth of information characterizing both chemical composition and morphology. The parallel nature of chemical imaging data makes it possible to analyze multiple samples simultaneously for applications that require high throughput analysis in addition to characterizing a single sample.

Principles

Chemical imaging shares the fundamentals of vibrational spectroscopic techniques, but provides additional information by way of the simultaneous acquisition of spatially resolved spectra. It combines the advantages of digital imaging with the attributes of spectroscopic measurements. Briefly, vibrational spectroscopy measures the interaction of light with matter. Photons that interact with a sample are either absorbed or scattered; photons of specific energy are absorbed, and the pattern of absorption provides information, or a fingerprint, on the molecules that are present in the sample.

On the other hand, in terms of the observation setup, chemical imaging can be carried out in one of the following modes: (optical) absorption, emission (fluorescence), (optical) transmission or scattering (Raman). A consensus currently exists that the fluorescence (emission) and Raman scattering modes are the most sensitive and powerful, but also the most expensive.

In a transmission measurement, the radiation goes through a sample and is measured by a detector placed on the far side of the sample. The energy transferred from the incoming radiation to the molecule(s) can be calculated as the difference between the quantity of photons that were emitted by the source and the quantity that is measured by the detector. In a diffuse reflectance measurement, the same energy difference measurement is made, but the source and detector are located on the same side of the sample, and the photons that are measured have re-emerged from the illuminated side of the sample rather than passed through it. The energy may be measured at one or multiple wavelengths; when a series of measurements are made, the response curve is called a spectrum.

A key element in acquiring spectra is that the radiation must somehow be energy selected – either before or after interacting with the sample. Wavelength selection can be accomplished with a fixed filter, tunable filter, spectrograph, an interferometer, or other devices. For a fixed filter approach, it is not efficient to collect a significant number of wavelengths, and multispectral data are usually collected. Interferometer-based chemical imaging requires that entire spectral ranges be collected, and therefore results in hyperspectral data. Tunable filters have the flexibility to provide either multi- or hyperspectral data, depending on analytical requirements.

Spectra may be measured one point at a time using a single element detector (single-point mapping), as a line-image using a linear array detector (typically 16 to 28 pixels) (linear array mapping), or as a two-dimensional image using a Focal Plane Array (FPA)(typically 256 to 16,384 pixels) (FPA imaging). For single-point the sample is moved in the x and y directions point-by-point using a computer-controlled stage. With linear array mapping, the sample is moved line-by-line with a computer-controlled stage. FPA imaging data are collected with a two-dimensional FPA detector, hence capturing the full desired field-of-view at one time for each individual wavelength, without having to move the sample. FPA imaging, with its ability to collect tens of thousands of spectra simultaneously is orders of magnitude faster than linear arrays which can typically collect 16 to 28 spectra simultaneously, which are in turn much faster than single-point mapping.

Terminology

Some words common in spectroscopy, optical microscopy and photography have been adapted or their scope modified for their use in chemical imaging. They include: resolution, field of view and magnification. There are two types of resolution in chemical imaging. The spectral resolution refers to the ability to resolve small energy differences; it applies to the spectral axis. The spatial resolution is the minimum distance between two objects that is required for them to be detected as distinct objects. The spatial resolution is influenced by the field of view, a physical measure of the size of the area probed by the analysis. In imaging, the field of view is a product of the magnification and the number of pixels in the detector array. The magnification is a ratio of the physical area of the detector array divided by the area of the sample field of view. Higher magnifications for the same detector image a smaller area of the sample.

Types of vibrational chemical imaging instruments

Chemical imaging has been implemented for mid-infrared, near-infrared spectroscopy and Raman spectroscopy. As with their bulk spectroscopy counterparts, each imaging technique has particular strengths and weaknesses, and are best suited to fulfill different needs.

Mid-infrared chemical imaging

Mid-infrared (MIR) spectroscopy probes fundamental molecular vibrations, which arise in the spectral range 2,500-25,000 nm. Commercial imaging implementations in the MIR region typically employ Fourier Transform Infrared (FT-IR) interferometers and the range is more commonly presented in wavenumber, $4,000 - 400 \text{ cm}^{-1}$. The MIR absorption bands tend to be relatively narrow and well-resolved; direct spectral interpretation is often possible by an experienced spectroscopist. MIR spectroscopy can distinguish subtle changes in chemistry and structure, and is often used for the identification of unknown materials. The absorptions in this spectral range are relatively strong; for this reason, sample presentation is important to limit the amount of material interacting with the incoming radiation in the MIR region. Most data collected in this range is collected in transmission mode through thin sections (~ 10 micrometres) of material. Water is a very strong absorber of MIR radiation and wet samples often require advanced sampling procedures (such as attenuated total reflectance). Commercial instruments include point and line mapping, and imaging. All employ an FT-IR interferometer as wavelength selective element and light source.

For types of MIR microscope, see [Microscopy#Infrared microscopy](#).

Atmospheric windows in the infrared spectrum are also employed to perform chemical imaging remotely. In these spectral regions the atmospheric gases (mainly water and CO_2) present low absorption and allow infrared viewing over kilometer distances. Target molecules can then be viewed using the selective absorption/emission processes described above. An example of the chemical imaging of a simultaneous release of SF_6 and NH_3 is shown in the image.



Remote chemical imaging of a simultaneous release of SF_6 and NH_3 at 1.5km using the FIRST imaging spectrometer^[24]

Near-infrared chemical imaging

The analytical near infrared (NIR) region spans the range from approximately 700-2,500 nm. The absorption bands seen in this spectral range arise from overtones and combination bands of O-H, N-H, C-H and S-H stretching and bending vibrations. Absorption is one to two orders of magnitude smaller in the NIR compared to the MIR; this phenomenon eliminates the need for extensive sample preparation. Thick and thin samples can be analyzed without any sample preparation, it is possible to acquire NIR chemical images through some packaging materials, and the technique can be used to examine hydrated samples, within limits. Intact samples can be imaged in transmittance or diffuse reflectance.

The lineshapes for overtone and combination bands tend to be much broader and more overlapped than for the fundamental bands seen in the MIR. Often, multivariate methods are used to separate spectral signatures of sample components. NIR chemical imaging is particularly useful for performing rapid, reproducible and non-destructive analyses of known materials^{[22] [23]}. NIR imaging instruments are typically based on one of two platforms: imaging using a tunable filter and broad band illumination, and line mapping employing an FT-IR interferometer as the wavelength filter and light source.

Raman chemical imaging

The Raman shift chemical imaging spectral range spans from approximately 50 to 4,000 cm^{-1} ; the actual spectral range over which a particular Raman measurement is made is a function of the laser excitation frequency. The basic principle behind Raman spectroscopy differs from the MIR and NIR in that the x-axis of the Raman spectrum is measured as a function of energy shift (in cm^{-1}) relative to the frequency of the laser used as the source of radiation. Briefly, the Raman spectrum arises from inelastic scattering of incident photons, which requires a change in polarizability with vibration, as opposed to infrared absorption, which requires a change in dipole moment with vibration. The end result is spectral information that is similar and in many cases complementary to the MIR. The Raman effect is weak - only about one in 10^7 photons incident to the sample undergoes Raman scattering. Both organic and inorganic materials possess a Raman spectrum; they generally produce sharp bands that are chemically specific. Fluorescence is a competing phenomenon and, depending on the sample, can overwhelm the Raman signal, for both bulk spectroscopy and imaging implementations.

Raman chemical imaging requires little or no sample preparation. However, physical sample sectioning may be used to expose the surface of interest, with care taken to obtain a surface that is as flat as possible. The conditions required for a particular measurement dictate the level of invasiveness of the technique, and samples that are sensitive to high power laser radiation may be damaged during analysis. It is relatively insensitive to the presence of water in the sample and is therefore useful for imaging samples that contain water such as biological material.

Fluorescence imaging (visible and NIR)

This emission microspectroscopy mode is the most sensitive in both visible and FT-NIR microspectroscopy, and has therefore numerous biomedical, biotechnological and agricultural applications. There are several powerful, highly specific and sensitive fluorescence techniques that are currently in use, or still being developed; among the former are FLIM, FRAP, FRET and FLIM-FRET; among the latter are NIR fluorescence and probe-sensitivity enhanced NIR fluorescence microspectroscopy and nanospectroscopy techniques (see "Further reading" section).

Sampling and samples

The value of imaging lies in the ability to resolve spatial heterogeneities in solid-state or gel/gel-like samples. Imaging a liquid or even a suspension has limited use as constant sample motion serves to average spatial information, unless ultra-fast recording techniques are employed as in fluorescence correlation microspectroscopy or FLIM observations where a single molecule may be monitored at extremely high (photon) detection speed. High-throughput experiments (such as imaging multi-well plates) of liquid samples can however provide valuable

information. In this case, the parallel acquisition of thousands of spectra can be used to compare differences between samples, rather than the more common implementation of exploring spatial heterogeneity within a single sample.

Similarly, there is no benefit in imaging a truly homogeneous sample, as a single point spectrometer will generate the same spectral information. Of course the definition of homogeneity is dependent on the spatial resolution of the imaging system employed. For MIR imaging, where wavelengths span from 3-10 micrometres, objects on the order of 5 micrometres may theoretically be resolved. The sampled areas are limited by current experimental implementations because illumination is provided by the interferometer. Raman imaging may be able to resolve particles less than 1 micrometre in size, but the sample area that can be illuminated is severely limited. With Raman imaging, it is considered impractical to image large areas and, consequently, large samples. FT-NIR chemical/hyperspectral imaging usually resolves only larger objects (>10 micrometres), and is better suited for large samples because illumination sources are readily available. However, FT-NIR microspectroscopy was recently reported to be capable of about 1.2 micron (micrometer) resolution in biological samples^[24] Furthermore, two-photon excitation FCS experiments were reported to have attained 15 nanometer resolution on biomembrane thin films with a special coincidence photon-counting setup.

Detection limit

The concept of the detection limit for chemical imaging is quite different than for bulk spectroscopy, as it depends on the sample itself. Because a bulk spectrum represents an average of the materials present, the spectral signatures of trace components are simply overwhelmed by dilution. In imaging however, each pixel has a corresponding spectrum. If the physical size of the trace contaminant is on the order of the pixel size imaged on the sample, its spectral signature will likely be detectable. If however, the trace component is dispersed homogeneously (relative to pixel image size) throughout a sample, it will not be detectable. Therefore, detection limits of chemical imaging techniques are strongly influenced by particle size, the chemical and spatial heterogeneity of the sample, and the spatial resolution of the image.

Data analysis

Data analysis methods for chemical imaging data sets typically employ mathematical algorithms common to single point spectroscopy or to image analysis. The reasoning is that the spectrum acquired by each detector is equivalent to a single point spectrum; therefore pre-processing, chemometrics and pattern recognition techniques are utilized with the similar goal to separate chemical and physical effects and perform a qualitative or quantitative characterization of individual sample components. In the spatial dimension, each chemical image is equivalent to a digital image and standard image analysis and robust statistical analysis can be used for feature extraction.

See also

- Multispectral image
- Microspectroscopy
- Imaging spectroscopy

Further reading

1. E. N. Lewis, P. J. Treado, I. W. Levin, Near-Infrared and Raman Spectroscopic Imaging, American Laboratory, 06/1994:16 (1994)
2. E. N. Lewis, P. J. Treado, R. C. Reeder, G. M. Story, A. E. Dowrey, C. Marcott, I. W. Levin, FTIR spectroscopic imaging using an infrared focal-plane array detector, Analytical Chemistry, 67:3377 (1995)
3. P. Colarusso, L. H. Kidder, I. W. Levin, J. C. Fraser, E. N. Lewis Infrared Spectroscopic Imaging: from Planetary to Cellular Systems, Applied Spectroscopy, 52 (3):106A (1998)

4. P. J. Treado I. W. Levin, E. N. Lewis, Near-Infrared Spectroscopic Imaging Microscopy of Biological Materials Using an Infrared Focal-Plane Array and an Acousto-Optic Tunable Filter (AOTF), *Applied Spectroscopy*, 48:5 (1994)
 5. Hammond, S.V., Clarke, F. C., Near-infrared microspectroscopy. In: *Handbook of Vibrational Spectroscopy*, Vol. 2, J.M. Chalmers and P.R. Griffiths Eds. John Wiley and Sons, West Sussex, UK, 2002, p.1405-1418
 6. L.H. Kidder, A.S. Haka, E.N. Lewis, Instrumentation for FT-IR Imaging. In: *Handbook of Vibrational Spectroscopy*, Vol. 2, J.M. Chalmers and P.R. Griffiths Eds. John Wiley and Sons, West Sussex, UK, 2002, pp.1386-1404
 7. J. Zhang; A. O'Connor; J. F. Turner II, Cosine Histogram Analysis for Spectral Image Data Classification, *Applied Spectroscopy*, Volume 58, Number 11, November 2004, pp. 1318-1324(7)
 8. J. F. Turner II; J. Zhang; A. O'Connor, A Spectral Identity Mapper for Chemical Image Analysis, *Applied Spectroscopy*, Volume 58, Number 11, November 2004, pp. 1308-1317(10)
 9. H. R. MORRIS, J. F. TURNER II, B. MUNRO, R. A. RYNTZ, P. J. TREADO, Chemical imaging of thermoplastic olefin (TPO) surface architecture, *Langmuir*, 1999, vol. 15, no8, pp. 2961-2972
 10. J. F. Turner II, Chemical imaging and spectroscopy using tunable filters: Instrumentation, methodology, and multivariate analysis, Thesis (PhD). UNIVERSITY OF PITTSBURGH, Source DAI-B 59/09, p. 4782, Mar 1999, 286 pages.
 11. P. Schwille.(2001). in *Fluorescence Correlation Spectroscopy. Theory and applications*. R. Rigler & E.S. Elson, eds., p. 360. Springer Verlag: Berlin.
 12. Schwille P., Oehlenschläger F. and Walter N. (1996). Analysis of RNA-DNA hybridization kinetics by fluorescence correlation spectroscopy, *Biochemistry* **35**:10182.
 13. FLIM | Fluorescence Lifetime Imaging Microscopy: Fluorescence, fluorophore chemical imaging, confocal emission microspectroscopy, FRET, cross-correlation fluorescence microspectroscopy ^[25].
 14. FLIM Applications: ^[25] "FLIM is able to discriminate between fluorescence emanating from different fluorophores and autofluorescing molecules in a specimen, even if their emission spectra are similar. It is, therefore, ideal for identifying fluorophores in multi-label studies. FLIM can also be used to measure intracellular ion concentrations without extensive calibration procedures (for example, Calcium Green) and to obtain information about the local environment of a fluorophore based on changes in its lifetime." FLIM is also often used in microspectroscopic/chemical imaging, or microscopic, studies to monitor spatial and temporal protein-protein interactions, properties of membranes and interactions with nucleic acids in living cells.
 15. Gadella TW Jr., *FRET and FLIM techniques*, 33. Imprint: Elsevier, ISBN 978-0-08-054958-3. (2008) 560 pages
 16. Langel FD, et al., Multiple protein domains mediate interaction between Bcl10 and Malt1, *J. Biol. Chem.*, (2008) 283(47):32419-31
 17. Clayton AH. , The polarized AB plot for the frequency-domain analysis and representation of fluorophore rotation and resonance energy homotransfer. *J Microscopy*. (2008) 232(2):306-12
 18. Clayton AH, et al., Predominance of activated EGFR higher-order oligomers on the cell surface. *Growth Factors* (2008) 20:1
 19. Plowman et al., Electrostatic Interactions Positively Regulate K-Ras Nanocluster Formation and Function. *Molecular and Cellular Biology* (2008) 4377–4385
 20. Belanis L, et al., Galectin-1 Is a Novel Structural Component and a Major Regulator of H-Ras Nanoclusters. *Molecular Biology of the Cell* (2008) 19:1404–1414
 21. Van Manen HJ, Refractive index sensing of green fluorescent proteins in living cells using fluorescence lifetime imaging microscopy. *Biophys J*. (2008) 94(8):L67-9
 22. Van der Krogt GNM, et al., A Comparison of Donor-Acceptor Pairs for Genetically Encoded FRET Sensors: Application to the Epac cAMP Sensor as an Example, *PLoS ONE*, (2008) 3(4):e1916
 23. Dai X, et al., Fluorescence intensity and lifetime imaging of free and micellar-encapsulated doxorubicin in living cells. *Nanomedicine*. (2008) **4**(1):49-56.
-

External links

- NIR Chemical Imaging in Pharmaceutical Industry ^[14]
- Pharmaceutical Process Analytical Technology: ^[15]
- NIR Chemical Imaging for Counterfeit Pharmaceutical Product Analysis ^[26]
- Chemical Imaging: Potential New Crime Busting Tool ^[27]
- Applications of Chemical Imaging in Research ^[28]

References

- [1] [http://www.imaging.net/chemical-imaging/Chemical imaging](http://www.imaging.net/chemical-imaging/Chemical%20imaging)
- [2] http://www.malvern.com/LabEng/products/sdi/bibliography/sdi_bibliography.htm E. N. Lewis, E. Lee and L. H. Kidder, Combining Imaging and Spectroscopy: Solving Problems with Near-Infrared Chemical Imaging. *Microscopy Today*, Volume 12, No. 6, 11/2004.
- [3] C.L. Evans and X.S. Xie.2008. Coherent Anti-Stokes Raman Scattering Microscopy: Chemical Imaging for Biology and Medicine., doi:10.1146/annurev.anchem.1.031207.112754 *Annual Review of Analytical Chemistry*, **1**: 883-909.
- [4] Diaspro, A., and Robello, M. (1999). Multi-photon Excitation Microscopy to Study Biosystems. *European Microscopy and Analysis.*, 5:5-7.
- [5] D.S. Mantus and G. H. Morrison. 1991. Chemical imaging in biology and medicine using ion microscopy., *Microchimica Acta*, **104**, (1-6) January 1991, doi: 10.1007/BF01245536
- [6] Bagatolli, L.A., and Gratton, E. (2000). Two-photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures. *Biophys J.*, 78:290-305.
- [7] Schwille, P., Haupts, U., Maiti, S., and Webb. W.(1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophysical Journal*, 77(10):2251-2265.
- [8] I.Lee, S. C. et al., (2001). One Micrometer Resolution NMR Microscopy. *J. Magn. Res.*, 150: 207-213.
- [9] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
- [10] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann and S. S. Korban, q-bio/0407006 (July 2004) (<http://arxiv.org/abs/q-bio/0407006>)
- [11] J. Dubois, G. Sando, E. N. Lewis, Near-Infrared Chemical Imaging, A Valuable Tool for the Pharmaceutical Industry, G.I.T. Laboratory Journal Europe, No. 1-2, 2007.
- [12] Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
- [13] Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology.(June 2004) I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin q-bio/0406047 (<http://arxiv.org/abs/q-bio/0406047>)
- [14] http://www.spectroscopyeurope.com/NIR_14_3.pdf
- [15] <http://www.fda.gov/cder/OPS/PAT.htm>
- [16] Eigen, M., and Rigler, R. (1994). Sorting single molecules: Applications to diagnostics and evolutionary biotechnology. *Proc. Natl. Acad. Sci. USA* 91:5740.
- [17] Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.
- [18] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy.2004.I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004) (<http://arxiv.org/abs/q-bio/0407006>)
- [19] Oehlenschläger F., Schwille P. and Eigen M. (1996). Detection of HIV-1 RNA by nucleic acid sequence-based amplification combined with fluorescence correlation spectroscopy, *Proc. Natl. Acad. Sci. USA* **93**:1281.
- [20] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
- [21] M. Chamberland, V. Farley, A. Vallières, L. Belhumeur, A. Villemare, J. Giroux et J. Legault, High-Performance Field-Portable Imaging Radiometric Spectrometer Technology For Hyperspectral imaging Applications, *Proc. SPIE* 5994, 59940N, September 2005.
- [22] Novel Techniques for Microspectroscopy and Chemical Imaging Analysis of Soybean Seeds and Embryos.(2002). Baianu, I.C., Costescu, D.M., and You, T. *Soy2002 Conference*, Urbana, Illinois.
- [23] Near Infrared Microspectroscopy, Chemical Imaging and NMR Analysis of Oil in Developing and Mutagenized Soybean Embryos in Culture. (2003). Baianu, I.C., Costescu, D.M., Hofmann, N., and Korban, S.S. *AOCS Meeting, Analytical Division*.
- [24] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
- [25] <http://www.nikoninstruments.com/infocenter.php?n=FLIM>
- [26] <http://www.spectroscopymag.com/spectroscopy/Near-IR+Spectroscopy/NIR-Chemical-Imaging-for-Counterfeit-Pharmaceutical/ArticleStandard/Article/detail/406629>

[27] <http://www.sciencedaily.com/releases/2007/08/070802103435.htm>

[28] <http://www3.imperial.ac.uk/vibrationalspectroscopyandchemicalimaging/research>

Spin polarization

Spin polarization is the degree to which the spin, i.e. the intrinsic angular momentum of elementary particles, is aligned with a given direction^[1]. This property may pertain to the spin, hence to the magnetic moment, of conduction electrons in ferromagnetic metals, such as iron, giving rise to spin polarized currents. It may also pertain to beams of particles, produced for particular aims, such as polarized neutron scattering or muon spin spectroscopy. Spin polarization of electrons or of nuclei, often called simply magnetization, is also produced by the application of a magnetic field, thanks to the Curie law and it is used to produce an induction signal in Electron spin resonance (ESR or EPR) and in Nuclear magnetic resonance (NMR).

Spin polarization is also important for spintronics, a branch of electronics. Magnetic semiconductors are being researched as possible spintronics materials.

The spin of free electrons is measured either by a LEED image from a clean wolfram-crystal (SPLEED)^{[2] [3] [4]} or by an electron microscope composed purely of electrostatic lenses and a gold foil as a sample. Back scattered electrons are decelerated by annular optics and focused onto a ring shaped electron mulitplier at about 15°. The position on the ring is recorded. This whole device is called a Mott-detector. Depending on their spin the electrons have the chance to hit the ring at different positions. 1% of the electrons are scattered in the foil. Of these 1% are collected by the detector and then about 30% of the electrons hit the detector at the wrong position. Both devices work due to spin orbit coupling.

References

- [1] J. Kessler (1976). *Polarized Electrons*. Springer Verlag Berlin Heidelberg. pp. 7–19.
 - [2] J. Kirschner and R. Feder (1979). "Spin Polarization in Double Diffraction of Low-Energy Electrons from W(001): Experiment and Theory". *Physical Review Letters* **42**: 1008–1011.
 - [3] M. Kalisvaart, M. R. O'Neill, T. W. Riddle, F. B. Dunning, and G. K. Walters (1977). "Electron-spin polarization in low-energy electron diffraction from tungsten (001)". *Physical Review B* **17**: 1570–1578.
 - [4] R. Feder (1976). "Spin Polarization in Low-Energy Electron Diffraction from W(001)". *Physical Review Letters* **36**: 598–600.
-

Polarized Neutron Spectroscopy

Triple-axis spectrometry (TAS, T also resolved as "three", S also resolved as "spectroscopy") is a technique used in inelastic neutron scattering. The instrument is referred to as triple-axis spectrometer (also called TAS). It allows measurement of the scattering function at any point in energy and momentum space physically accessible by the spectrometer.

History

The triple-axis spectrometry method was first developed by Bertram Brockhouse at the NRX research reactor at the Chalk River Laboratories in Canada. The first results from the prototype triple-axis spectrometer were published in January 1955 and the first true triple-axis spectrometer was built in 1956. Bertram Brockhouse shared the 1994 Nobel prize for Physics for this development, which allowed elementary excitations, such as phonons and magnons, to be observed directly. The Nobel citation was "for pioneering contributions to the development of neutron scattering techniques for studies of condensed matter" and "for the development of neutron spectroscopy".

TAS Instruments in current use

FRM-II Forschungsneutronenquelle Heinz Maier-Leibnitz

- PANDA ^[1] - a cold neutron triple-axis spectrometer.
- PUMA ^[2] - a thermal neutron triple-axis spectrometer with multianalyser-detector option.

Institut Laue-Langevin

- IN1 ^[3] - a hot neutron triple-axis spectrometer.
- IN3 ^[4] - a thermal neutron triple-axis spectrometer for tests.
- IN8 ^[5] - a high-flux thermal neutron triple-axis spectrometer.
- IN12 ^[6] - a cold neutron triple-axis spectrometer.
- IN14 ^[7] - a cold neutron triple-axis spectrometer with polarized neutron capability.
- IN20 ^[8] - a thermal neutron triple-axis spectrometer with polarized neutron capability.
- IN22 ^[9] - a thermal neutron triple-axis spectrometer with polarized neutron capability.
- D10 ^[10] - a thermal neutron four-circle diffractometer with a triple-axis energy analysis option.

CEA/Saclay Laboratoire Léon Brillouin

- 1T-1 ^[11] - a double-focusing thermal neutron triple-axis spectrometer.
- 2T-1 ^[12] - a thermal neutron triple-axis spectrometer.
- 4F-1 ^[13] - a cold neutron triple-axis spectrometer.
- 4F-2 ^[14] - a cold neutron triple-axis spectrometer.

NIST Center for Neutron Research

- SPINS ^[15] - a cold neutron triple-axis spectrometer with polarized neutron capability.
 - BT-7 ^[16] - a thermal neutron triple-axis spectrometer with polarized neutron capability.
 - BT-9 ^[16] - a thermal neutron triple-axis spectrometer.
-

MURR^[17] University of Missouri Research Reactor

- Triax^[18] - a thermal neutron triple-axis spectrometer.

External links

- Nobelprize.org page for the 1994 Nobel Prize for Physics^[19]

References

- [1] <http://wwwnew.frm2.tum.de/en/science/instruments/spectrometers/panda.html>
- [2] <http://wwwnew.frm2.tum.de/en/science/instruments/spectrometers/puma.html>
- [3] <http://www.ill.fr/YellowBook/IN1/>
- [4] <http://www.ill.fr/YellowBook/IN3/>
- [5] <http://www.ill.fr/YellowBook/IN8/>
- [6] <http://www.ill.fr/YellowBook/IN12/>
- [7] <http://www.ill.fr/YellowBook/IN14/>
- [8] <http://www.ill.fr/YellowBook/IN20/>
- [9] <http://www.ill.fr/YellowBook/IN22/>
- [10] <http://www.ill.fr/YellowBook/D10/>
- [11] <http://www-llb.cea.fr/spectros/pdf/1t1-llb.pdf>
- [12] <http://www-llb.cea.fr/spectros/pdf/2t1-llb.pdf>
- [13] <http://www-llb.cea.fr/spectros/pdf/4f1-llb.pdf>
- [14] <http://www-llb.cea.fr/spectros/pdf/4f2-llb.pdf>
- [15] <http://www.ncnr.nist.gov/instruments/spins/>
- [16] http://www.ncnr.nist.gov/instruments/bt7_new/
- [17] <http://www.murr.missouri.edu/>
- [18] http://www.murr.missouri.edu/rd_material_sciences_instrumentation_triax.php
- [19] http://nobelprize.org/nobel_prizes/physics/laureates/1994/

Polarized Muon Spectroscopy

Muon spin spectroscopy is an experimental technique based on the implantation of spin polarized muons in matter and on the detection of the influence of the atomic, molecular or crystalline surroundings on their spin motion. The motion of the muon spin is due to the magnetic field experienced by the particle and may provide information on its local environment in a very similar way to other magnetic resonance techniques, such as electron spin resonance (ESR or EPR) and, more closely, nuclear magnetic resonance (NMR).

Acronym

In analogy with the acronyms for these previously established spectroscopies, the muon spin spectroscopy is also known as μ SR, which stands for **muon spin rotation**, or relaxation, or resonance, depending respectively on whether the muon spin motion is predominantly a rotation (more precisely a precession around a still magnetic field), or a relaxation towards an equilibrium direction, or, again, a more complex dynamics dictated by the addition of short radio frequency pulses.

How it works

The time scale on which the spin motion may be exploited is that of the muon decay, i.e. a few mean lifetimes, each roughly $2.2\ \mu\text{s}$ (2.2 millionths of a second). Both the production of muon beams with nearly perfect alignment of the spin to the beam direction (what was referred to above as spin polarization and caused by the spontaneous symmetry breaking), and the ability to detect the muon spin direction at the instant of the muon decay rely on the violation of parity, which takes place whenever weak forces are at play.

In short this means that certain elementary events happen only when including clockwise (or only when including counterclockwise) rotations. For instance, the positive muon decays into a positron plus two neutrinos and the positron is preferentially emitted in the direction of the muon spin. Therefore it would most often *see* the spin as a counterclockwise rotation while flying away from the decay point.

Spin alignment allows the production of a muon beam with an aligned magnetic moment. Muons are injected into the material under investigation as short-lived *spies* [1] sending information from the interior back out to the experimental apparatus. These muons are able to send a message from inside the crystal about the local magnetic field in their surroundings. After some time (mean lifetime $2.2\ \mu\text{s}$) these spies decay and emit positrons. A beam of aligned muons produces asymmetric positron radiation. The asymmetry of positron radiation contains information about the direction of local magnetic field in the moment of muon decay. Taking into consideration the initial direction of muon magnetic moment and the time interval between the moment of injection and moment of muon decay we can calculate the precession frequency (how rapidly the muon's magnetic moment rotates). The frequency of magnetic moment precession depends on the local magnetic field. Larmor precession is appeared with z -direction magnetic field and only decay in $2.2\ \mu\text{s}$. But when x -direction magnetic field is applied in muon, the rate of decay is enhanced by gaussian with depolarization rate.

Since 1987 this method was used to measure internal magnetic fields inside high-temperature superconductors. High-temperature superconductors are Type II superconductors, in which the local magnetic fields inside the superconductor depend on the superconducting carrier density—one of the significant parameters of any superconductor (see for example the Bardeen–Cooper–Shrieffer theory of superconductors).

Applications

Muon Spin Rotation and Relaxation are mostly performed with positive muons. They are well suited to the study of magnetic fields at the atomic scale inside matter, such as those produced by various kinds of magnetism and/or superconductivity encountered in compounds occurring in nature or artificially produced by modern material science.

The London penetration depth is one of the most important parameters characterizing a superconductor because its inverse square provides a measure of the density n_s of Cooper pairs. The dependence of n_s on temperature and magnetic field directly indicates the symmetry of the superconducting gap. Muon spin spectroscopy provides a way to measure the penetration depth, and so has been used to study high-temperature cuprate superconductors since their discovery in 1986.

Other important fields of application of μ SR exploit the fact that positive muons capture electrons to form muonium atoms which behave chemically as light isotopes of the hydrogen atom. This allows investigation of the largest known "isotope effect" in some of the simplest types of chemical reactions, as well as the early stages of formation of radicals in organic chemicals. Muonium is also studied as an analogue of hydrogen in semiconductors, where hydrogen is one of the most ubiquitous impurities.

Facilities

μ SR requires a particle accelerator for the production of a muon beam. This is presently achieved at few large scale facilities in the world: the CMMS[2] continuous source at TRIUMF in Vancouver, Canada; the LMU [3] continuous source at the Paul Scherrer Institut (PSI) in Villigen, Switzerland; the ISIS and RIKEN-RAL [4] pulsed sources at the Rutherford Appleton Laboratory in Chilton, United Kingdom; and the J-PARC facility in Tokai, Japan, where a new pulsed source is being built to replace that at KEK in Tsukuba, Japan. Muon beams are also available at the Laboratory of Nuclear Problems, Joint Institute for Nuclear Research (JINR) in Dubna, Russia. The International Society for μ SR Spectroscopy (ISMS [5]) exists to promote the worldwide advancement of μ SR. Membership in the society is open free of charge to all individuals in academia, government laboratories and industry who have an interest in the society's goals.

See also

- Muon
- Muonium
- Nuclear magnetic resonance

References

- ISIS Introductory course in μ SR [6]
 - introduction to μ SR [7]
 - μ SR Brochure [8] (a 3.2 MB PDF file)
 - CMMS [2]: TRIUMF Center for Molecular and Materials Science
 - ISIS [9]
-

References

- [1] <http://musr.org/intro/musr/>
- [2] <http://musr.org/>
- [3] <http://lmu.web.psi.ch/>
- [4] <http://nectar.nd.rl.ac.uk/~rikenral/index.html>
- [5] <http://musr.org/isms/>
- [6] http://www.isis.rl.ac.uk/muons/index.htm?content_area=/muons/trainingcourse/course2005.htm&side_nav=/muons/muonsSideNav.htm
- [7] <http://musr.org/intro/musr/intro.htm>
- [8] <http://muon.neutron-eu.net/muon/files/muSRBrochure.pdf>
- [9] <http://www.isis.stfc.ac.uk/>

Time-resolved spectroscopy

In physics and physical chemistry, **time-resolved spectroscopy** is the study of dynamic processes in materials or chemical compounds by means of spectroscopic techniques. Most often, processes are studied that occur after illumination of a material, but in principle, the technique can be applied to any process which leads to a change in properties of a material. With the help of pulsed lasers, it is possible to study processes which occur on time scales as short as 10^{-14} seconds. The rest of the article discusses different types of time-resolved spectroscopy.

Transient-absorption spectroscopy

Transient-absorption spectroscopy is an extension of absorption spectroscopy. Here, the absorbance at a particular wavelength or range of wavelengths of a sample is measured as a function of time after excitation by a flash of light. In a typical experiment, both the light for excitation ('pump') and the light for measuring the absorbance ('probe') are generated by a pulsed laser. If the process under study is slow, then the time resolution can be obtained with a continuous (i.e., not pulsed) probe beam and repeated conventional spectrophotometric techniques.

Examples of processes that can be studied:

- Optical gain spectroscopy of semiconductor laser materials.
- Chemical reactions that are initiated by light (or 'photoinduced chemical reactions');
- The transfer of excitation energy between molecules, parts of molecules, or molecules and their environment;
- The behaviour of electrons that are freed from a molecule or crystalline material.

Other multiple-pulse techniques

Transient spectroscopy as discussed above is a technique that involves two pulses. There are many more techniques that employ two or more pulses, such as:

- Photon echoes.
- Four-wave mixing (involves three laser pulses)

The interpretation of experimental data from these techniques is usually much more complicated than in transient-absorption spectroscopy.

Nuclear magnetic resonance and electron spin resonance are often implemented with multiple-pulse techniques, though with radio waves and micro waves instead of visible light.

Time-resolved infrared spectroscopy

Time-resolved infrared (TRIR) spectroscopy also employs a two-pulse, "pump-probe" methodology. The pump pulse is typically in the UV region and is often generated by a high-powered Nd:YAG laser whilst the probe beam is in the infrared region. This technique currently operates down to the picosecond time regime and surpasses transient absorption and emission spectroscopy by providing *structural* information on the excited-state kinetics of both dark and emissive states.

Time-resolved fluorescence spectroscopy

Time-resolved fluorescence spectroscopy is an extension of fluorescence spectroscopy. Here, the fluorescence of a sample is monitored as a function of time after excitation by a flash of light. The time resolution can be obtained in a number of ways, depending on the required sensitivity and time resolution:

- With fast detection electronics (nanoseconds and slower);
- With a streak camera (picoseconds and slower);
- With optical gating (femtoseconds-nanoseconds) - a short laser pulse acts as a gate for the detection of fluorescence light; only fluorescence light that arrives at the detector at the same time as the gate pulse is detected. This technique has the best time resolution, but the efficiency is rather low. An extension of this optical gating technique is to use a "Kerr gate", which allows the scattered Raman signal to be collected before the (slower) fluorescence signal overwhelms it. This technique can greatly improve the signal:noise ratio of Raman spectra.

Terahertz spectroscopy

Terahertz frequency radiation for spectroscopy is typically generated in one of three ways:

- time domain terahertz spectroscopy (TDTS), using ultrashort laser pulses
 - photomixing, mixing two radiation sources to generate their difference frequency
 - Fourier transform spectroscopy, using a blackbody radiation source
-

Applied spectroscopy

Applied spectroscopy is the application of various spectroscopic methods for detection and identification of different elements/compounds in solving problems in the fields of forensics, medicine, oil industry, atmospheric chemistry, pharmacology, etc.

Spectroscopic methods

Among the more common spectroscopic methods used for analysis is FTIR spectroscopy, where chemical bonds can be detected through their characteristic infra-red absorption frequencies or wavelengths. UV spectroscopy is used where strong absorption of ultra-violet radiation occurs in a substance. Such groups are known as chromophores and include aromatic groups, conjugated system of bonds, carbonyl groups and so on. NMR spectroscopy detects hydrogen atoms in specific environments, and complements both IR and UV spectroscopy. The use of Raman spectroscopy is growing for more specialist applications.

There are also derivative methods such as infrared microscopy which allows very small areas to be analysed in an optical microscope.

One method of elemental analysis which is important in forensic analysis is EDX performed in the environmental scanning electron microscope, or ESEM. The method involves analysis of back-scattered X-rays from the sample as a result of interaction with the electron beam.

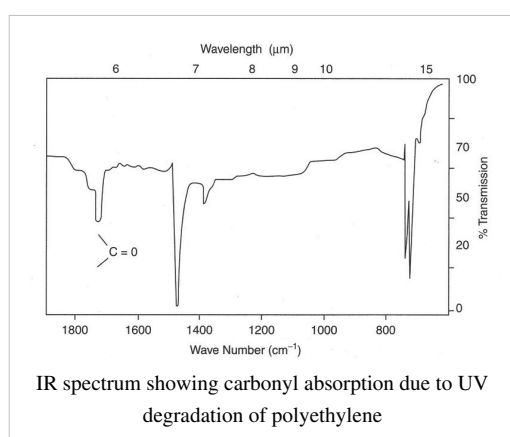
Sample preparation

In all three spectroscopic methods, the sample usually needs to be present in solution, which may present problems during forensic examination because it necessarily involves sampling solid from the object to be examined.

FTIR: Three types of samples can be analyzed, a solution (KBr), a powder, or a film. A solid film is the easiest and most straight forward sample type to test.

Analysis of polymers

Many polymer degradation mechanisms can be followed using infra-red spectroscopy, such as UV degradation and oxidation, amongst many other failure modes.



UV degradation

Many polymers are attacked by UV radiation at vulnerable points in their chain structures. Thus polypropylene suffers severe cracking in sunlight unless anti-oxidants are added. The point of attack occurs at the tertiary carbon atom present in every repeat unit, causing oxidation and finally chain breakage. Polyethylene is also susceptible to UV degradation, especially those variants which are branched polymers such as LDPE. The branch points are tertiary carbon atoms, so polymer degradation starts there and results in chain cleavage, and embrittlement. In the example shown at left, carbonyl groups were readily detected by IR

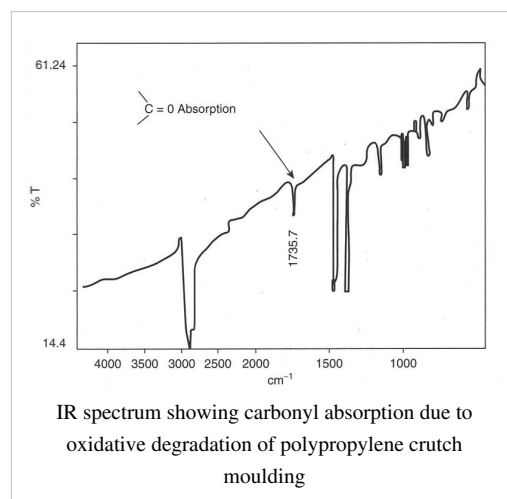
spectroscopy from a cast thin film. The product was a road cone which had cracked in service, and many similar cones also failed because an anti-UV additive had not been used.

Oxidation

Polymers are susceptible to attack by atmospheric oxygen, especially at elevated temperatures encountered during processing to shape. Many process methods such as extrusion and injection moulding involve pumping molten polymer into tools, and the high temperatures needed for melting may result in oxidation unless precautions are taken. For example, a forearm crutch suddenly snapped and the user was severely injured in the resulting fall. The crutch had fractured across a polypropylene insert within the aluminium tube of the device, and infra-red spectroscopy of the material showed that it had oxidised, possible as a result of poor moulding.

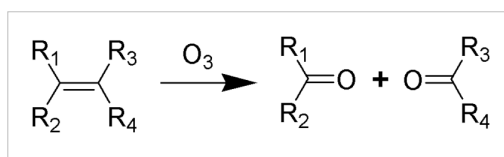
Oxidation is usually relatively easy to detect owing to the strong absorption by the carbonyl group in the spectrum of polyolefins.

Polypropylene has a relatively simple spectrum with few peaks at the carbonyl position (like polyethylene). Oxidation tends to start at tertiary carbon atoms because free radicals here are more stable, so last longer and are attacked by oxygen. The carbonyl group can be further oxidised to break the chain, so weakening the material by lowering the molecular weight, and cracks start to grow in the regions affected.

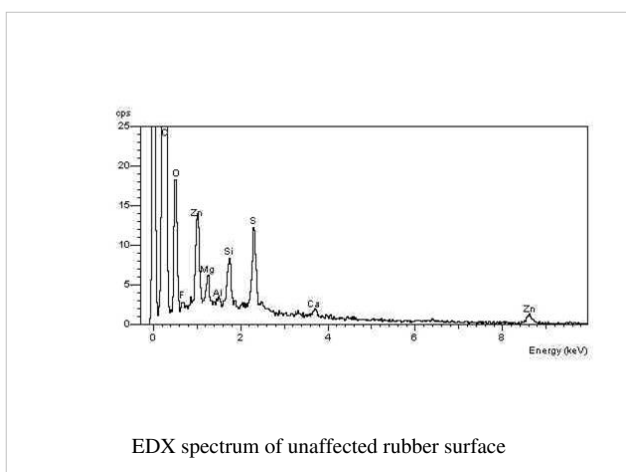
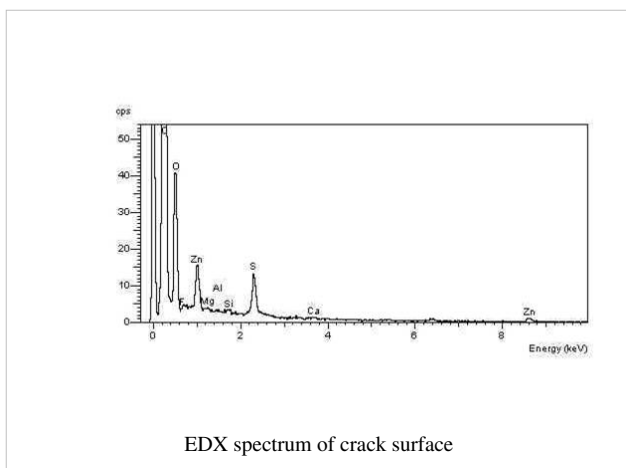


Ozonolysis

The reaction occurring between double bonds and ozone is known as ozonolysis when one molecule of the gas reacts with the double bond:



The immediate result is formation of an ozonide, which then decomposes rapidly so that the double bond is cleaved. This is the critical step in chain breakage when polymers are attacked. The strength of polymers depends on the chain molecular weight or degree of polymerization, the higher the chain length, the greater the mechanical strength (such as tensile strength). By cleaving the chain, the molecular weight drops rapidly and there comes a point when it has little strength whatsoever, and a crack forms. Further attack occurs in the freshly exposed crack surfaces and the crack grows steadily until it completes a circuit and the product separates or fails. In the case of a seal or a tube, failure occurs when the wall of the device is penetrated.



The carbonyl end groups which are formed are usually aldehydes or ketones, which can oxidise further to carboxylic acids. The net result is a high concentration of elemental oxygen on the crack surfaces, which can be detected using Energy-dispersive X-ray spectroscopy in the environmental SEM, or ESEM. The spectrum at left shows the high oxygen peak compared with a constant sulphur peak. The spectrum at right shows the unaffected elastomer surface spectrum, with a relatively low oxygen peak compared with the sulphur peak. The spectra were obtained during an investigation into ozone cracking of diaphragm seals in a semi-conductor fabrication factory.

See also

- Absorption spectroscopy
- Infrared spectroscopy correlation table
- Infrared spectroscopy
- Forensic chemistry
- Forensic engineering
- Forensic polymer engineering
- Polymer degradation
- Polymer engineering
- Spectroscopy

- Society for Applied Spectroscopy

References

- *Forensic Materials Engineering: Case Studies* by Peter Rhys Lewis, Colin Gagg, Ken Reynolds, CRC Press (2004).
- Peter R Lewis and Sarah Hainsworth, *Fuel Line Failure from stress corrosion cracking*, Engineering Failure Analysis, 13 (2006) 946-962.

External links

- Museum of failed products ^[1]
- New Forensic course ^[2]
- The journal Engineering Failure Analysis ^[3]
- Forensic science and engineering ^[4]

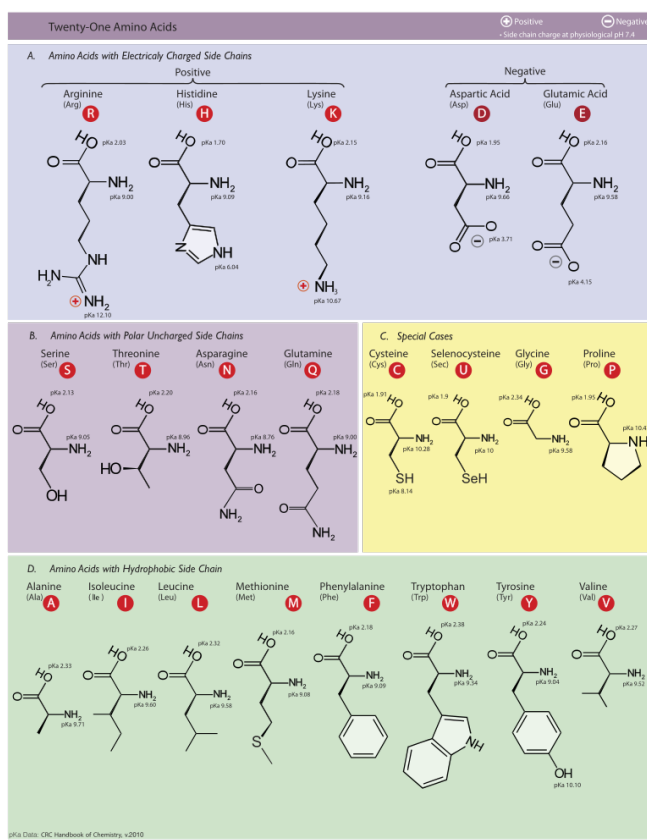
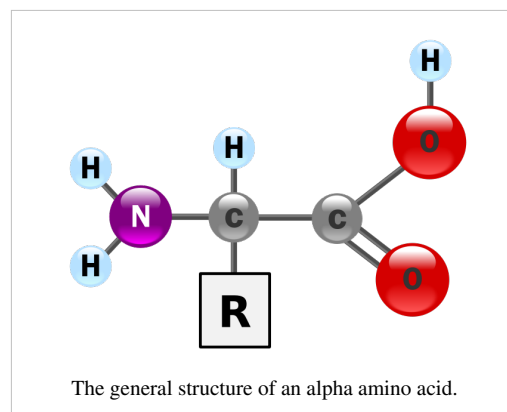
References

- [1] <http://materials.open.ac.uk/mem/index.htm>
- [2] <http://openlearn.open.ac.uk/file.php/2980/formats/print.htm>
- [3] http://www.elsevier.com/wps/find/journaldescription.cws_home/30190/description#description
- [4] <http://www.forensic-courses.com/wordpress/?p=42;>

Amino acids

Amino acids are molecules containing an amine group, a carboxylic acid group and a side chain that varies between different amino acids. These molecules contain the key elements of Carbon, Hydrogen, Oxygen, and Nitrogen. These molecules are particularly important in biochemistry, where this term refers to alpha-amino acids with the general formula $\text{H}_2\text{NCH(R)COOH}$, where R is an organic substituent.^[1] In an alpha amino acid, the amino and carboxylate groups are attached to the same carbon atom, which is called the α -carbon. The various alpha amino acids differ in which side chain (R group) is attached to their alpha carbon. These side chains can vary in size from just a hydrogen atom in glycine, to a methyl group in alanine, through to a large heterocyclic group in tryptophan.

Amino acids are critical to life, and have many functions in metabolism. One particularly important function is as the building blocks of proteins, which are linear chains of amino acids. Every protein is chemically defined by this primary structure, its unique sequence of amino acid residues, which in turn define the three-dimensional structure of the protein. Just as the letters of the alphabet can be combined to form an almost endless variety of words, amino acids can be linked together in varying sequences to form a vast variety of proteins.^[2] Amino acids are also important in many other biological molecules, such as forming parts of coenzymes, as in S-adenosylmethionine, or as precursors for the biosynthesis of molecules such as heme. Due to this central role in biochemistry, amino acids are very important in



The twenty-one amino acids found in eukaryotes, grouped according to their side chains' pKa's and charge at physiological pH 7.4

nutrition. Amino acids are commonly used in food technology and industry. For example, monosodium glutamate is a common flavor enhancer that gives foods the taste called *umami*. They are also used in industry. Applications include the production of biodegradable plastics, drugs and chiral catalysts.

History

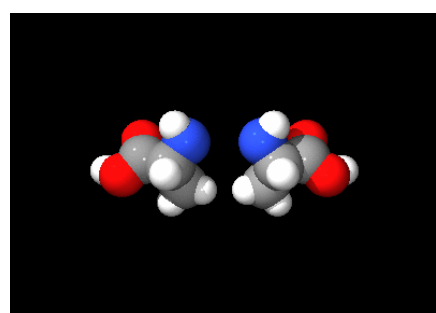
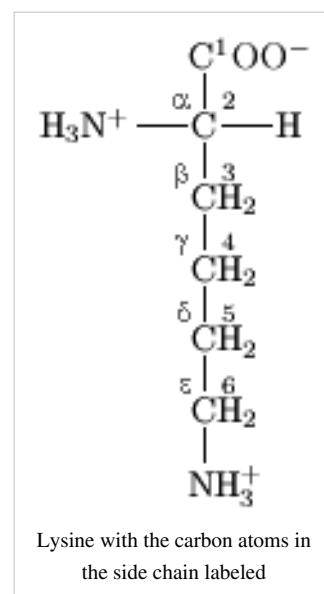
The first few amino acids were discovered in the early 1800s. In 1806, the French chemists Louis-Nicolas Vauquelin and Pierre Jean Robiquet isolated a compound in asparagus that proved to be asparagine, the first amino acid to be discovered.^[3] ^[4] Another amino acid that was discovered in the early 19th century was cystine, in 1810,^[5] although its monomer, cysteine, was discovered much later, in 1884.^[4] ^[6] Glycine and leucine were also discovered around this time, in 1820.^[7]

General structure

In the structure shown at the top of the page, **R** represents a side chain specific to each amino acid. The carbon atom next to the carbonyl group is called the α -carbon and amino acids with a side chain bonded to this carbon are referred to as *alpha amino acids*. These are the most common form found in nature. In the alpha amino acids, the α -carbon is a chiral carbon atom, with the exception of glycine.^[8] In amino acids that have a carbon chain attached to the α -carbon (such as lysine, shown to the right) the carbons are labeled in order as α , β , γ , δ , and so on.^[9] In some amino acids, the amine group is attached to the β or γ -carbon, and these are therefore referred to as *beta* or *gamma amino acids*.

Amino acids are usually classified by the properties of their side chain into four groups. The side chain can make an amino acid a weak acid or a weak base, and a hydrophile if the side chain is polar or a hydrophobe if it is nonpolar.^[8] The chemical structures of the twenty-two standard amino acids, along with their chemical properties, are described more fully in the article on these proteinogenic amino acids.

The phrase "branched-chain amino acids" or BCAA refers to the amino acids having aliphatic side chains that are non-linear; these are leucine, isoleucine, and valine. Proline is the only proteinogenic amino acid whose side group links to the α -amino group and, thus, is also the only proteinogenic amino acid containing a secondary amine at this position.^[8] Chemically, proline is therefore an imino acid since it lacks a primary amino group,^[10] although it is still classed as an amino acid in the current biochemical nomenclature,^[11] and may also be called an "N-alkylated alpha-amino acid".^[12]



The two optical isomers of alanine, D-Alanine and L-Alanine

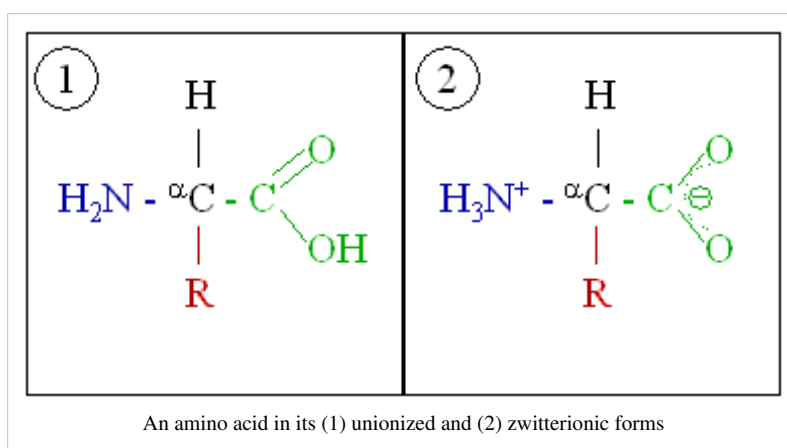
Isomerism

Of the standard α -amino acids, all but glycine can exist in either of two optical isomers, called L or D amino acids, which are mirror images of each other (*see also Chirality*). While L-amino acids represent all of the amino acids found in proteins during translation in the ribosome, D-amino acids are found in some proteins produced by enzyme posttranslational modification after translation and translocation to the endoplasmic reticulum, as in exotic sea-dwelling organisms such as cone snails.^[13] They are also abundant components of the

peptidoglycan cell walls of bacteria.^[14] and D-serine may act as a neurotransmitter in the brain.^[15] The L and D convention for amino acid configuration refers not to the optical activity of the amino acid itself, but rather to the optical activity of the isomer of glyceraldehyde from which that amino acid can theoretically be synthesized (D-glyceraldehyde is dextrorotary; L-glyceraldehyde is levorotary). Alternatively, the (*S*) and (*R*) designators are used to indicate the absolute stereochemistry. Almost all of the amino acids in proteins are (*S*) at the α carbon, with cysteine being (*R*) and glycine non-chiral.^[16] Cysteine is unusual since it has a sulfur atom at the first position in its side-chain, which has a larger atomic mass than the groups attached to the α -carbon in the other standard amino acids, thus the (*R*) instead of (*S*).

Zwitterions

Amino acids have both amine and carboxylic acid functional groups and are therefore both an acid and a base at the same time.^[8] At a certain pH known as the isoelectric point an amino acid has no overall charge, since the number of protonated ammonium groups (positive charges) and deprotonated carboxylate groups (negative charges) are equal.^[17] The amino acids all have

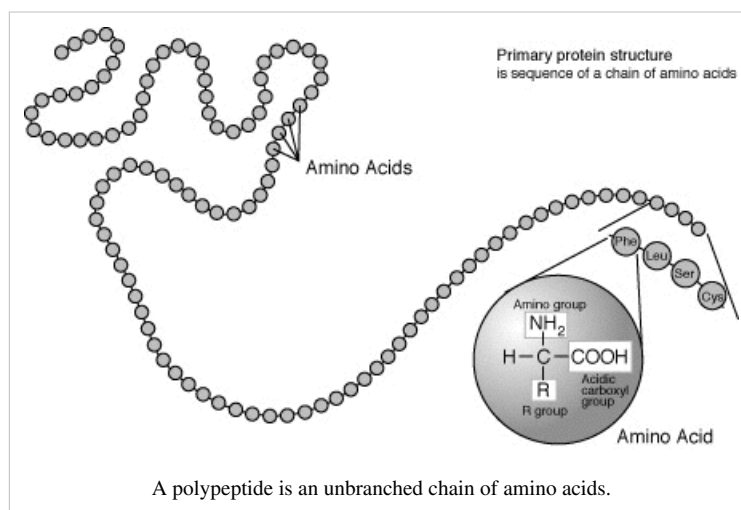


different isoelectric points. The ions produced at the isoelectric point have both positive and negative charges and are known as a *zwitterion*, which comes from the German word *Zwitter* meaning "hermaphrodite" or "hybrid".^[18] Amino acids can exist as zwitterions in solids and in polar solutions such as water, but not in the gas phase.^[19] Zwitterions have minimal solubility at their isoelectric point and an amino acid can be isolated by precipitating it from water by adjusting the pH to its particular isoelectric point.

Occurrence and functions in biochemistry

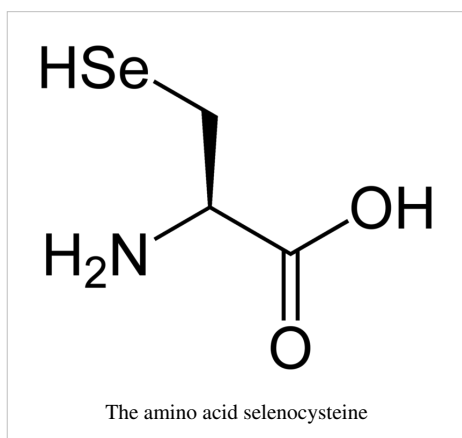
Standard amino acids

Amino acids are the structural units that make up proteins. They join together to form short polymer chains called peptides or longer chains called either polypeptides or proteins. These polymers are linear and unbranched, with each amino acid within the chain attached to two neighbouring amino acids. The process of making proteins is called *translation* and involves the step-by-step addition of amino acids to a growing protein chain by a ribozyme that is called a ribosome.^[20] The order in which the amino acids are added is read through the genetic code from an mRNA template, which is a RNA copy of one of the organism's genes



Twenty-two amino acids are naturally incorporated into polypeptides and are called proteinogenic or standard amino acids.^[8] Of these twenty-two, twenty are directly encoded by the universal genetic code. The remaining two,

selenocysteine and pyrrolysine, are incorporated into proteins by unique synthetic mechanisms. Selenocysteine is incorporated when the mRNA being translated includes a SECIS element, which causes the UGA codon to encode selenocysteine instead of a stop codon.^[21] Pyrrolysine is used by some methanogenic archaea in enzymes that they use to produce methane. It is coded for with the codon UAG, which is normally a stop codon in other organisms.^[22]

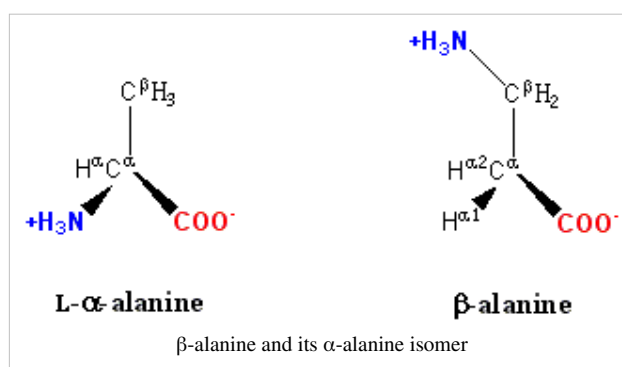


Non-standard amino acids

Aside from the twenty-two standard amino acids, there are a vast number of "non-standard" amino acids. These non-standard amino acids found in proteins are formed by post-translational modification, which is modification after translation in protein synthesis. These modifications are often essential for the function or regulation of a protein; for example, the carboxylation of glutamate allows for better binding of calcium cations,^[23] and the hydroxylation of proline is critical for maintaining connective tissues.^[24] Another example is the formation of hypusine in the translation initiation factor EIF5A, through modification of a lysine residue.^[25] Such modifications can

also determine the localization of the protein, e.g., the addition of long hydrophobic groups can cause a protein to bind to a phospholipid membrane.^[26]

Examples of nonstandard amino acids that are not found in proteins include lanthionine, 2-aminoisobutyric acid, dehydroalanine and the neurotransmitter gamma-aminobutyric acid. Nonstandard amino acids often occur as intermediates in the metabolic pathways for standard amino acids — for example ornithine and citrulline occur in the urea cycle, part of amino acid catabolism (see below).^[27] A rare exception to the dominance of α -amino acids in biology is the β -amino acid beta alanine (3-aminopropanoic acid), which is used in plants and microorganisms in the synthesis of pantothenic acid (vitamin B5), a component of coenzyme A.^[28]



In human nutrition

When taken up into the human body from the diet, the twenty two standard amino acids are either used to synthesize proteins and other biomolecules or oxidized to urea and carbon dioxide as a source of energy.^[29] The oxidation pathway starts with the removal of the amino group by a transaminase, the amino group is then fed into the urea cycle. The other product of transamination is a keto acid that enters the citric acid cycle.^[30] Glucogenic amino acids can also be converted into glucose, through gluconeogenesis.^[31]

Pyrrolysine trait is restricted to several microbes, and only one organism has both Pyl and Sec. Of the twenty-two standard amino acids, eight are called essential amino acids because the human body cannot synthesize them from other compounds at the level needed for normal growth, so they must be obtained from food.^[32] However, the situation is quite complicated since cysteine, taurine, tyrosine, histidine and arginine are semiessential amino acids in children, because the metabolic pathways that synthesize these amino acids are not fully developed.^[33] ^[34] The amounts required also depend on the age and health of the individual, so it is hard to make general statements about the dietary requirement for some amino acids.

Essential	Nonessential
Isoleucine	Alanine
Leucine	Asparagine
Lysine	Aspartic Acid
Methionine	Cysteine*
Phenylalanine	Glutamic Acid
Threonine	Glutamine*
Tryptophan	Glycine*
Valine	Proline*
	Selenocysteine*
	Serine*
	Tyrosine*
	Arginine*
	Histidine*

(*) Essential only in certain cases.^{[35] [36]}

Non-protein functions

In humans, non-protein amino acids also have important roles as metabolic intermediates, such as in the biosynthesis of the neurotransmitter gamma-aminobutyric acid. Many amino acids are used to synthesize other molecules, for example:

- Tryptophan is a precursor of the neurotransmitter serotonin.^[37]
- Glycine is a precursor of porphyrins such as heme.^[38]
- Arginine is a precursor of nitric oxide.^[39]
- Ornithine and S-adenosylmethionine are precursors of polyamines.^[40]
- Aspartate, glycine and glutamine are precursors of nucleotides.^[41]
- Phenylalanine is a precursor of various phenylpropanoids which are important in plant metabolism.

However, not all of the functions of other abundant non-standard amino acids are known, for example taurine is a major amino acid in muscle and brain tissues, but although many functions have been proposed, its precise role in the body has not been determined.^[42]

Some non-standard amino acids are used as defenses against herbivores in plants.^[43] For example canavanine is an analogue of arginine that is found in many legumes,^[44] and in particularly large amounts in *Canavalia gladiata* (sword bean).^[45] This amino acid protects the plants from predators such as insects and can cause illness in people if some types of legumes are eaten without processing.^[46] The non-protein amino acid mimosine is found in other species of legume, particularly *Leucaena leucocephala*.^[47] This compound is an analogue of tyrosine and can poison animals that graze on these plants.

Uses in technology

Amino acids are used for a variety of applications in industry but their main use is as additives to animal feed. This is necessary since many of the bulk components of these feeds, such as soybeans, either have low levels or lack some of the essential amino acids: lysine, methionine, threonine, and tryptophan are most important in the production of these feeds.^[48] The food industry is also a major consumer of amino acids, particularly glutamic acid, which is used as a flavor enhancer,^[49] and Aspartame (aspartyl-phenylalanine-1-methyl ester) as a low-calorie artificial

sweetener.^[50] The remaining production of amino acids is used in the synthesis of drugs and cosmetics.^[48]

Amino acid derivative	Pharmaceutical application
5-HTP (5-hydroxytryptophan)	Experimental treatment for depression. ^[51]
L-DOPA (L-dihydroxyphenylalanine)	Treatment for Parkinsonism. ^[52]
Eflornithine	Drug that inhibits ornithine decarboxylase and is used in the treatment of sleeping sickness. ^[53]

Chemical building blocks

Amino acids are important as low-cost feedstocks. These compounds are used in chiral pool synthesis as enantiomerically-pure building blocks.^[54]

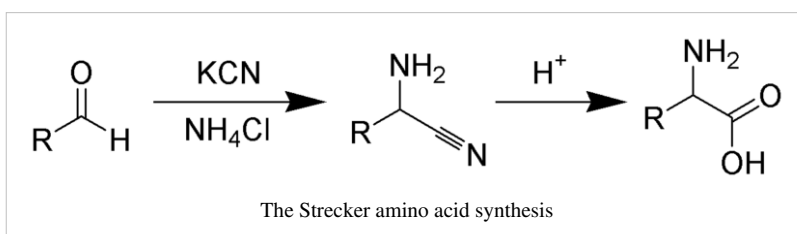
Amino acids have been investigated as precursors chiral catalysts, e.g. for asymmetric hydrogenation reactions, although no commercial applications exist.^[55]

Biodegradable plastics

Amino acids are under development as components of a range of biodegradable polymers. These materials have applications as environmentally-friendly packaging and in medicine in drug delivery and the construction of prosthetic implants. These polymers include polypeptides, polyamides, polyesters, polysulfides and polyurethanes with amino acids either forming part of their main chains or bonded as side chains. These modifications alter the physical properties and reactivities of the polymers.^[56] An interesting example of such materials is polyaspartate, a water-soluble biodegradable polymer that may have applications in disposable diapers and agriculture.^[57] Due to its solubility and ability to chelate metal ions, polyaspartate is also being used as a biodegradable anti-scaling agent and a corrosion inhibitor.^[58] ^[59] In addition, the aromatic amino acid tyrosine is being developed as a possible replacement for toxic phenols such as bisphenol A in the manufacture of polycarbonates.^[60]

Reactions

As amino acids have both a primary amine group and a primary carboxyl group, these chemicals can undergo most of the reactions associated with these functional groups. These include nucleophilic addition, amide bond formation and imine formation for the amine group and esterification, amide bond formation and decarboxylation for the carboxylic acid group.^[61] The multiple side chains of amino acids can also undergo chemical reactions.^[62] The types of these reactions are determined by the groups on these side chains and are therefore different between the various types of amino acid.



Chemical synthesis

Several methods exist to synthesize amino acids. One of the oldest methods, begins with the bromination at the α -carbon of a carboxylic acid.

Nucleophilic substitution with

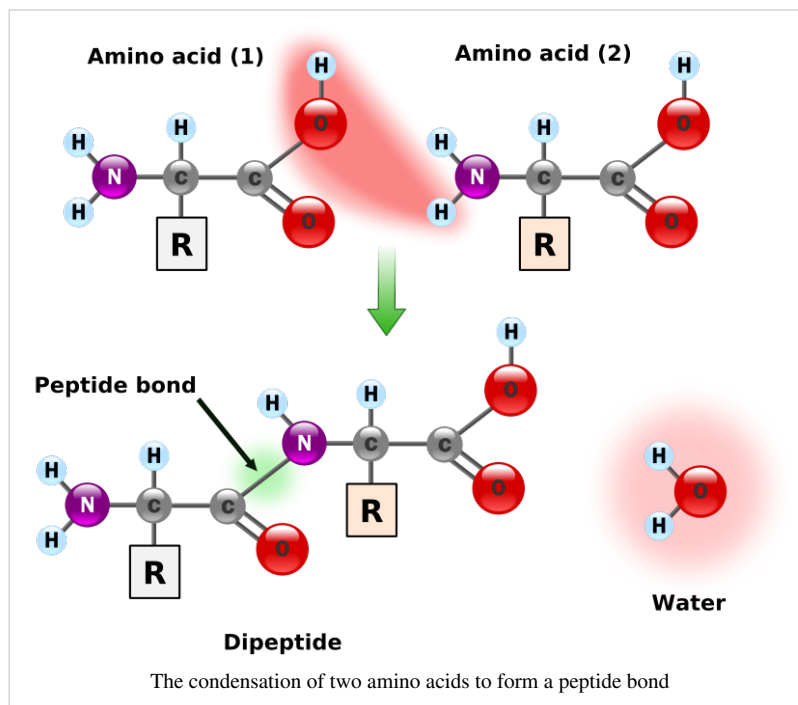
ammonia then converts the alkyl bromide to the amino acid.^[63] Alternatively, the Strecker amino acid synthesis involves the treatment of an aldehyde with potassium cyanide and ammonia, this produces an α -amino nitrile as an intermediate. Hydrolysis of the nitrile in acid then yields a α -amino acid.^[64] Using ammonia or ammonium salts in this reaction gives unsubstituted amino acids, while substituting primary and secondary amines will yield substituted amino acids.^[65] Likewise, using ketones, instead of aldehydes, gives α,α -disubstituted amino acids.^[66] The classical synthesis gives racemic mixtures of α -amino acids as products, but several alternative procedures using asymmetric

auxiliaries^[67] or asymmetric catalysts^{[68] [69]} have been developed.^[70]

Currently the most adopted method is an automated synthesis on a solid support (e.g. polystyrene beads), using protecting groups (e.g. Fmoc and t-Boc) and activating groups (e.g. DCC and DIC).

Peptide bond formation

As both the amine and carboxylic acid groups of amino acids can react to form amide bonds, one amino acid molecule can react with another and become joined through an amide linkage. This polymerization of amino acids is what creates proteins. This condensation reaction yields the newly formed peptide bond and a molecule of water. In cells, this reaction does not occur directly; instead the amino acid is first activated by attachment to a transfer RNA molecule through an ester bond. This aminoacyl-tRNA is produced in an ATP-dependent reaction carried out by an aminoacyl tRNA synthetase.^[71] This aminoacyl-tRNA is then a substrate for the ribosome, which catalyzes the



attack of the amino group of the elongating protein chain on the ester bond.^[72] As a result of this mechanism, all proteins made by ribosomes are synthesized starting at their N-terminus and moving towards their C-terminus.

However, not all peptide bonds are formed in this way. In a few cases, peptides are synthesized by specific enzymes. For example, the tripeptide glutathione is an essential part of the defenses of cells against oxidative stress. This peptide is synthesized in two steps from free amino acids.^[73] In the first step gamma-glutamylcysteine synthetase condenses cysteine and glutamic acid through a peptide bond formed between the side-chain carboxyl of the glutamate (the gamma carbon of this side chain) and the amino group of the cysteine. This dipeptide is then condensed with glycine by glutathione synthetase to form glutathione.^[74]

In chemistry, peptides are synthesized by a variety of reactions. One of the most used in solid-phase peptide synthesis, which uses the aromatic oxime derivatives of amino acids as activated units. These are added in sequence onto the growing peptide chain, which is attached to a solid resin support.^[75] The ability to easily synthesize vast numbers of different peptides by varying the types and order of amino acids (using combinatorial chemistry) has made peptide synthesis particularly important in creating libraries of peptides for use in drug discovery through high-throughput screening.^[76]

Biosynthesis and catabolism

In plants, nitrogen is first assimilated into organic compounds in the form of glutamate, formed from alpha-ketoglutarate and ammonia in the mitochondrion. In order to form other amino acids, the plant uses transaminases to move the amino group to another alpha-keto carboxylic acid. For example, aspartate aminotransferase converts glutamate and oxaloacetate to alpha-ketoglutarate and aspartate.^[77] Other organisms use transaminases for amino acid synthesis too. Transaminases are also involved in breaking down amino acids.

Degrading an amino acid often involves moving its amino group to alpha-ketoglutarate, forming glutamate. In many vertebrates, the amino group is then removed through the urea cycle and is excreted in the form of urea. However, amino acid degradation can produce uric acid or ammonia instead. For example, serine dehydratase converts serine to pyruvate and ammonia.^[78]

Nonstandard amino acids are usually formed through modifications to standard amino acids. For example, homocysteine is formed through the transsulfuration pathway or by the demethylation of methionine via the intermediate metabolite S-adenosyl methionine,^[42] while hydroxyproline is made by a posttranslational modification of proline.^[79]

Microorganisms and plants can synthesize many uncommon amino acids. For example, some microbes make 2-aminoisobutyric acid and lanthionine, which is a sulfide-bridged derivative of alanine. Both of these amino acids are found in peptidic lantibiotics such as alamethicin.^[80] While in plants, 1-aminocyclopropane-1-carboxylic acid is a small disubstituted cyclic amino acid that is a key intermediate in the production of the plant hormone ethylene.^[81]

Hydrophilic and hydrophobic amino acids

Depending on the polarity of the side chain, amino acids vary in their hydrophilic or hydrophobic character.^[8] These properties are important in protein structure and protein–protein interactions. The importance of the physical properties of the side chains comes from the influence this has on the amino acid residues' interactions with other structures, both within a single protein and between proteins. The distribution of hydrophilic and hydrophobic amino acids determines the tertiary structure of the protein, and their physical location on the outside structure of the proteins influences their quaternary structure.^[8]

For example, soluble proteins have surfaces rich with polar amino acids like serine and threonine, while integral membrane proteins tend to have outer rings of hydrophobic amino acids that anchor them into the lipid bilayer. In the case part-way between these two extremes, peripheral membrane proteins have a patch of hydrophobic amino acids on their surface that locks onto the membrane. Similarly, proteins that have to bind to positively-charged molecules have surfaces rich with negatively charged amino acids like glutamate and aspartate, while proteins binding to negatively-charged molecules have surfaces rich with positively charged chains like lysine and arginine. Recently a new scale of hydrophobicity based on the free energy of hydrophobic association has been proposed.^[82]

The hydrophilic and hydrophobic interactions of proteins are not always the result of the properties of their amino acid sidechains. This is because a range of posttranslational modifications can attach other chemical groups to the amino acids in proteins. For example, these modifications can produce hydrophobic lipoproteins,^[83] or hydrophilic glycoproteins.^[84] These type of modification allow the reversible targeting of a protein to a membrane. For example, the addition and removal of the fatty acid palmitic acid to cysteine residues in some signaling proteins causes the proteins to attach and then detach from cell membranes.^[85]

Table of standard amino acid abbreviations and side chain properties

Amino Acid	3-Letter ^[86]	1-Letter ^[86]	Side chain polarity ^[86]	Side chain charge (pH 7.4) ^[86]	Hydropathy index ^[87]	Absorbance λ_{max} (nm) ^[88]	ϵ at λ_{max} ($\times 10^{-3}$ M ⁻¹ cm ⁻¹) ^[88]
Alanine	Ala	A	nonpolar	neutral	1.8		
Arginine	Arg	R	polar	positive	-4.5		
Asparagine	Asn	N	polar	neutral	-3.5		
Aspartic acid	Asp	D	polar	negative	-3.5		
Cysteine	Cys	C	nonpolar	neutral	2.5	250	0.3
Glutamic acid	Glu	E	polar	negative	-3.5		
Glutamine	Gln	Q	polar	neutral	-3.5		
Glycine	Gly	G	nonpolar	neutral	-0.4		
Histidine	His	H	polar	positive(10%), neutral(90%)	-3.2	211	5.9
Isoleucine	Ile	I	nonpolar	neutral	4.5		
Leucine	Leu	L	nonpolar	neutral	3.8		
Lysine	Lys	K	polar	positive	-3.9		
Methionine	Met	M	nonpolar	neutral	1.9		
Phenylalanine	Phe	F	nonpolar	neutral	2.8	257, 206, 188	0.2, 9.3, 60.0
Proline	Pro	P	nonpolar	neutral	-1.6		
Serine	Ser	S	polar	neutral	-0.8		
Threonine	Thr	T	polar	neutral	-0.7		
Tryptophan	Trp	W	nonpolar	neutral	-0.9	280, 219	5.6, 47.0
Tyrosine	Tyr	Y	polar	neutral	-1.3	274, 222, 193	1.4, 8.0, 48.0
Valine	Val	V	nonpolar	neutral	4.2		

In addition to the specific amino acid codes, placeholders are used in cases where chemical or crystallographic analysis of a peptide or protein can not conclusively determine the identity of a residue.

Ambiguous Amino Acids	3-Letter	1-Letter
Asparagine or aspartic acid	Asx	B
Glutamine or glutamic acid	Glx	Z
Leucine or Isoleucine	Xle	J
Unspecified or unknown amino acid	Xaa	X

Unk is sometimes used instead of **Xaa**, but is less standard.

Additionally, many non-standard amino acids have a specific code. For example, several peptide drugs, such as Bortezomib or MG132 are artificially synthesized and retain their protecting groups, which have specific codes. Bortezomib is Pyz-Phe-boroLeu and MG132 is Z-Leu-Leu-Leu-al. Additionally, To aid in the analysis of protein structure, photocrosslinking amino acids are available. These include photoleucine (**pLeu**) and photomethionine (**pMet**).^[89]

See also

- Amino acid dating
- Beta amino acid
- Degron
- Glucogenic amino acid
- Homochirality
- Leucines
- Proteinogenic amino acid (including chemical structures)
- Table of codons, 3-nucleotide sequences that encode each amino acid

Further reading

- Doolittle, R.F. (1989) Redundancies in protein sequences. In *Predictions of Protein Structure and the Principles of Protein Conformation* (Fasman, G.D. ed) Plenum Press, New York, pp. 599–623
- David L. Nelson and Michael M. Cox, *Lehninger Principles of Biochemistry*, 3rd edition, 2000, Worth Publishers, ISBN 1-57259-153-6
- Meierhenrich, U.J.: *Amino acids and the asymmetry of life*, Springer-Verlag, Berlin, New York, 2008. ISBN 978-3-540-76885-2
- Morelli, Robert J. "Studies of amino acid absorption from the small intestine." San Francisco: Morelli, 1952.

External links

- Amino acids overview^[90] physical-chemistry properties, 3D structures, etc
- List of Standard Amino Acids^[91] The Detailed PDF List of Standard Amino Acids (including 3D depictions)
- Nomenclature and Symbolism for Amino Acids and Peptides^[92] IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN)
- Molecular Expressions: The Amino Acid Collection^[93] – Has detailed information and microscopy photographs of each amino acid.
- Amino acid properties^[94] – Properties of the amino acids (a tool aimed mostly at molecular geneticists trying to understand the meaning of mutations)
- Synthesis of Amino Acids and Derivatives^[95]
- Learn the 20 proteinogenic amino acids online^[96]
- The origin of the single-letter code for the amino acids^[97]
- Amino acid solution's pH, titration and isoelectric point calculation free spreadsheet^[98]

References

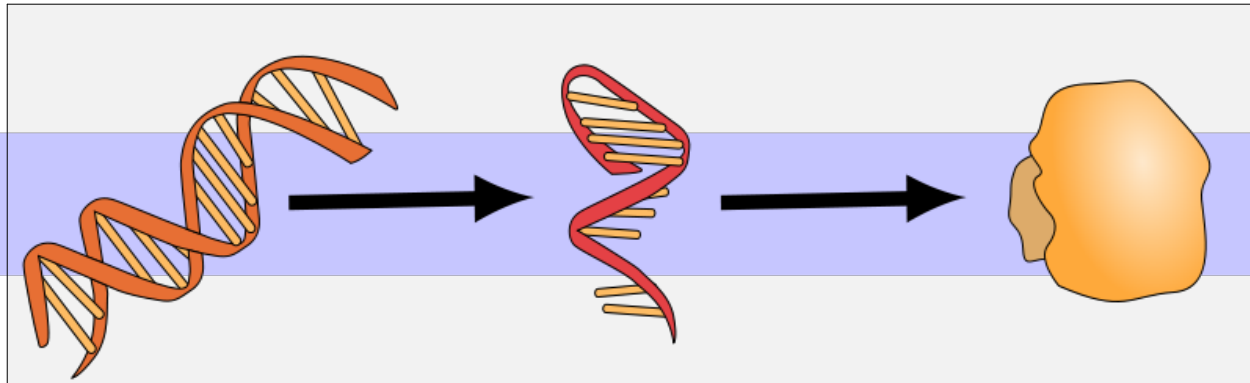
- [1] Proline is an exception to this general formula. It lacks the NH_2 group because of the cyclization of the side chain and it therefore falls under the category of special structured amino acids.
- [2] "The Structures of Life" (<http://publications.nigms.nih.gov/structlife/chapter1.html>). National Institute of General Medical Sciences. . Retrieved 2008-05-20.
- [3] Vauquelin LN, Robiquet PJ (1806). "The discovery of a new plant principle in *Asparagus sativus*". *Annales de Chimie* **57**: 88–93.
- [4] Anfinsen CB, Edsall JT, Richards FM (1972). *Advances in Protein Chemistry*. New York: Academic Press. pp. 99, 103. ISBN 978-0-12-034226-6.
- [5] Wollaston WH (1810). "On cystic oxide, a new species of urinary calculus". *Philosophical Transactions of the Royal Society of London* **100**: 223–30. doi:10.1098/rstl.1810.0015.
- [6] Baumann E (1884). "Über cystin und cystein". *Z Physiol Chemie* **8**: 299.
- [7] Braconnot HM (1820). "Sur la conversion des matières animales en nouvelles substances par le moyen de l'acide sulfurique". *Ann Chim Phys Ser 2* **13**: 113–25.
- [8] Creighton, Thomas H. (1993). "Chapter 1". *Proteins: structures and molecular properties*. San Francisco: W. H. Freeman. ISBN 978-0-7167-7030-5.

- [9] "Nomenclature and Symbolism for Amino Acids and Peptides" (<http://www.chem.qmul.ac.uk/iupac/AminoAcid/AA1n2.html>). IUPAC-IUB Joint Commission on Biochemical Nomenclature. 1983. . Retrieved 2008-11-17.
- [10] Jodidi, S. L. (1926-03-01). "The Formol Titration of Certain Amino Acids". *Journal of the American Chemical Society* **48** (3): 751–753. doi:10.1021/ja01414a033.
- [11] Liebecq, Claude, ed (1992). *Biochemical Nomenclature and Related Documents* (2nd ed.). Portland Press. pp. 39–69. ISBN 978-1-85578-005-7.
- [12] Smith, Anthony D. (1997). *Oxford dictionary of biochemistry and molecular biology*. Oxford: Oxford University Press. pp. 535. ISBN 978-0-19-854768-6. OCLC 37616711.
- [13] Pisarewicz K, Mora D, Pflueger FC, Fields GB, Marí F (May 2005). "Polypeptide chains containing D-gamma-hydroxyvaline". *Journal of the American Chemical Society* **127** (17): 6207–15. doi:10.1021/ja050088m. PMID 15853325.
- [14] van Heijenoort J (March 2001). "Formation of the glycan chains in the synthesis of bacterial peptidoglycan". *Glycobiology* **11** (3): 25R–36R. doi:10.1093/glycob/11.3.25R. PMID 11320055.
- [15] Wolosker H, Dumin E, Balan L, Foltyn VN (July 2008). "D-amino acids in the brain: D-serine in neurotransmission and neurodegeneration". *The FEBS Journal* **275** (14): 3514–26. doi:10.1111/j.1742-4658.2008.06515.x. PMID 18564180.
- [16] Hatem, Salama Mohamed Ali (2006). "Gas chromatographic determination of Amino Acid Enantiomers in tobacco and bottled wines" (<http://geb.uni-giessen.de/geb/volltexte/2006/3038/index.html>). University of Giessen. . Retrieved 2008-11-17.
- [17] Fennema OR (1996). *Food Chemistry* (3rd ed.). CRC Press. pp. 327–8. ISBN 0-8247-9691-8.
- [18] Simmons, William J.; Gerhard Meisenberg (2006). *Principles of medical biochemistry*. Mosby Elsevier. p. 19. ISBN 0-323-02942-6.
- [19] Remko M, Rode BM (February 2006). "Effect of metal ions (Li+, Na+, K+, Mg2+, Ca2+, Ni2+, Cu2+, and Zn2+) and water coordination on the structure of glycine and zwitterionic glycine". *The journal of physical chemistry. A* **110** (5): 1960–7. doi:10.1021/jp054119b. PMID 16451030.
- [20] Rodnina MV, Beringer M, Wintermeyer W (January 2007). "How ribosomes make peptide bonds". *Trends in Biochemical Sciences* **32** (1): 20–6. doi:10.1016/j.tibs.2006.11.007. PMID 17157507.
- [21] Driscoll DM, Copeland PR (2003). "Mechanism and regulation of selenoprotein synthesis". *Annual Review of Nutrition* **23**: 17–40. doi:10.1146/annurev.nutr.23.011702.073318. PMID 12524431.
- [22] Krzycki JA (December 2005). "The direct genetic encoding of pyrrolysine". *Current Opinion in Microbiology* **8** (6): 706–12. doi:10.1016/j.mib.2005.10.009. PMID 16256420.
- [23] Vermeer C (March 1990). "Gamma-carboxyglutamate-containing proteins and the vitamin K-dependent carboxylase" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1131186>). *The Biochemical Journal* **266** (3): 625–36. PMID 2183788. PMC 1131186.
- [24] Bhattacharjee A, Bansal M (March 2005). "Collagen structure: the Madras triple helix and the current scenario". *IUBMB Life* **57** (3): 161–72. doi:10.1080/15216540500090710. PMID 16036578.
- [25] Park MH (February 2006). "The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A)" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2494880>). *Journal of Biochemistry* **139** (2): 161–9. doi:10.1093/jb/mvj034. PMID 16452303. PMC 2494880.
- [26] Blenis J, Resh MD (December 1993). "Subcellular localization specified by protein acylation and phosphorylation". *Current Opinion in Cell Biology* **5** (6): 984–9. doi:10.1016/0955-0674(93)90081-Z. PMID 8129952.
- [27] Curis E, Nicolis I, Moinard C, *et al.* (November 2005). "Almost all about citrulline in mammals". *Amino Acids* **29** (3): 177–205. doi:10.1007/s00726-005-0235-4. PMID 16082501.
- [28] Coxon KM, Chakauya E, Ottenhof HH, *et al.* (August 2005). "Pantothenate biosynthesis in higher plants". *Biochemical Society Transactions* **33** (Pt 4): 743–6. doi:10.1042/BST0330743. PMID 16042590.
- [29] Sakami W, Harrington H (1963). "Amino acid metabolism". *Annual Review of Biochemistry* **32**: 355–98. doi:10.1146/annurev.bi.32.070163.002035. PMID 14144484.
- [30] Brosnan JT (April 2000). "Glutamate, at the interface between amino acid and carbohydrate metabolism" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=10736367>). *The Journal of Nutrition* **130** (4S Suppl): 988S–90S. PMID 10736367. .
- [31] Young VR, Ajami AM (September 2001). "Glutamine: the emperor or his clothes?" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=11533293>). *The Journal of Nutrition* **131** (9 Suppl): 2449S–59S; discussion 2486S–7S. PMID 11533293. .
- [32] Young VR (August 1994). "Adult amino acid requirements: the case for a major revision in current recommendations" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=8064412>). *The Journal of Nutrition* **124** (8 Suppl): 1517S–1523S. PMID 8064412. .
- [33] Imura K, Okada A (January 1998). "Amino acid metabolism in pediatric patients". *Nutrition* **14** (1): 143–8. doi:10.1016/S0899-9007(97)00230-X. PMID 9437700.
- [34] Lourenço R, Camilo ME (2002). "Taurine: a conditionally essential amino acid in humans? An overview in health and disease". *Nutrición Hospitalaria* **17** (6): 262–70. PMID 12514918.
- [35] Fürst P, Stehle P (June 2004). "What are the essential elements needed for the determination of amino acid requirements in humans?" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=15173430>). *The Journal of Nutrition* **134** (6 Suppl): 1558S–1565S. PMID 15173430. .
- [36] Reeds PJ (July 2000). "Dispensable and indispensable amino acids for humans" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=10867060>). *The Journal of Nutrition* **130** (7): 1835S–40S. PMID 10867060. .
- [37] Savelieva KV, Zhao S, Pogorelov VM, *et al.* (2008). "Genetic disruption of both tryptophan hydroxylase genes dramatically reduces serotonin and affects behavior in models sensitive to antidepressants" (<http://www.pubmedcentral.nih.gov/articlerender>).

- fcgi?tool=pmcentrez&artid=2565062). *PloS ONE* **3** (10): e3301. doi:10.1371/journal.pone.0003301. PMID 18923670. PMC 2565062.
- [38] Shemin D, Rittenberg D (1 December 1946). "The biological utilization of glycine for the synthesis of the protoporphyrin of hemoglobin" (<http://www.jbc.org/cgi/reprint/166/2/621>). *Journal of Biological Chemistry* **166** (2): 621–5. .
- [39] Tejero J, Biswas A, Wang ZQ, *et al.* (November 2008). "Stabilization and characterization of a heme-oxy reaction intermediate in inducible nitric-oxide synthase" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2586280>). *The Journal of Biological Chemistry* **283** (48): 33498–507. doi:10.1074/jbc.M806122200. PMID 18815130. PMC 2586280.
- [40] Rodríguez-Caso C, Montañez R, Cascante M, Sánchez-Jiménez F, Medina MA (August 2006). "Mathematical modeling of polyamine metabolism in mammals". *The Journal of Biological Chemistry* **281** (31): 21799–812. doi:10.1074/jbc.M602756200. PMID 16709566.
- [41] Stryer, Lubert; Berg, Jeremy Mark; Tymoczko, John L. (2002). *Biochemistry*. San Francisco: W.H. Freeman. pp. 693–8. ISBN 0-7167-4684-0.
- [42] Brosnan JT, Brosnan ME (June 2006). "The sulfur-containing amino acids: an overview" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=16702333>). *The Journal of Nutrition* **136** (6 Suppl): 1636S–1640S. PMID 16702333. .
- [43] Hylin, John W. (1969). "Toxic peptides and amino acids in foods and feeds". *Journal of Agricultural and Food Chemistry* **17** (3): 492–6. doi:10.1021/jf60163a003.
- [44] Turner, B. L.; Harborne, J. B. (1967). "Distribution of canavanine in the plant kingdom". *Phytochemistry* **6**: 863–66. doi:10.1016/S0031-9422(00)86033-1.
- [45] Ekanayake S, Skog K, Asp NG (May 2007). "Canavanine content in sword beans (*Canavalia gladiata*): analysis and effect of processing". *Food and Chemical Toxicology* **45** (5): 797–803. doi:10.1016/j.fct.2006.10.030. PMID 17187914.
- [46] Rosenthal GA (2001). "L-Canavanine: a higher plant insecticidal allelochemical". *Amino Acids* **21** (3): 319–30. doi:10.1007/s007260170017. PMID 11764412.
- [47] Hammond AC (May 1995). "Leucaena toxicosis and its control in ruminants" (<http://jas.fass.org/cgi/pmidlookup?view=long&pmid=7665380>). *Journal of Animal Science* **73** (5): 1487–92. PMID 7665380. .
- [48] Leuchtenberger W, Huthmacher K, Drauz K (November 2005). "Biotechnological production of amino acids and derivatives: current status and prospects". *Applied Microbiology and Biotechnology* **69** (1): 1–8. doi:10.1007/s00253-005-0155-y. PMID 16195792.
- [49] Garattini S (April 2000). "Glutamic acid, twenty years later" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=10736350>). *The Journal of Nutrition* **130** (4S Suppl): 901S–9S. PMID 10736350. .
- [50] Stegink LD (July 1987). "The aspartame story: a model for the clinical testing of a food additive" (<http://www.ajcn.org/cgi/pmidlookup?view=long&pmid=3300262>). *The American Journal of Clinical Nutrition* **46** (1 Suppl): 204–15. PMID 3300262. .
- [51] Turner EH, Loftis JM, Blackwell AD (March 2006). "Serotonin a la carte: supplementation with the serotonin precursor 5-hydroxytryptophan". *Pharmacology & Therapeutics* **109** (3): 325–38. doi:10.1016/j.pharmthera.2005.06.004. PMID 16023217.
- [52] Kostrzewa RM, Nowak P, Kostrzewa JP, Kostrzewa RA, Brus R (March 2005). "Peculiarities of L: -DOPA treatment of Parkinson's disease". *Amino Acids* **28** (2): 157–64. doi:10.1007/s00726-005-0162-4. PMID 15750845.
- [53] Heby O, Persson L, Rentala M (August 2007). "Targeting the polyamine biosynthetic enzymes: a promising approach to therapy of African sleeping sickness, Chagas' disease, and leishmaniasis". *Amino Acids* **33** (2): 359–66. doi:10.1007/s00726-007-0537-9. PMID 17610127.
- [54] Hanessian, S. (1993). "Reflections on the total synthesis of natural products: Art, craft, logic, and the chiron approach". *Pure and Applied Chemistry* **65**: 1189–204. doi:10.1351/pac199365061189.
- [55] Blaser, Hans Ulrich (1992). "The chiral pool as a source of enantioselective catalysts and auxiliaries". *Chemical Reviews* **92** (5): 935–52. doi:10.1021/cr00013a009.
- [56] Sanda, Fumio; Endo, Takeshi (1999). "Feature Article Syntheses and functions of polymers based on amino acids". *Macromolecular Chemistry and Physics* **200**: 2651–61. doi:10.1002/(SICI)1521-3935(19991201)200:12<2651::AID-MACP2651>3.0.CO;2-P.
- [57] Gross, R. A.; Kalra, B. (2002). "Biodegradable Polymers for the Environment" (<http://www.sciencemag.org/cgi/content/abstract/297/5582/803>). *Science* **297** (5582): 803–807. doi:10.1126/science.297.5582.803. PMID 12161646. .
- [58] Low, K. C.; Wheeler, A. P.; Koskan, L. P. (1996). *Commercial poly(aspartic acid) and Its Uses*. Advances in Chemistry Series. **248**. Washington, D.C.: American Chemical Society.
- [59] Thombre, S.M.; Sarwade, B.D. (2005). "Synthesis and Biodegradability of Polyaspartic Acid: A Critical Review" (<http://www.informaworld.com/index/718581646.pdf>). *Journal of Macromolecular Science, Part A* **42** (9): 1299–1315. doi:10.1080/10601320500189604. .
- [60] Bourke, S. L.; Kohn, J. (2003). "Polymers derived from the amino acid l-tyrosine: polycarbonates, polyarylates and copolymers with poly(ethylene glycol)" (<http://linkinghub.elsevier.com/retrieve/pii/S0169409X03000383>). *Advanced Drug Delivery Reviews* **55** (4): 447–466. doi:10.1016/S0169-409X(03)00038-3. PMID 12706045. .
- [61] Elmore, Donald Trevor; Barrett, G. C. (1998). *Amino acids and peptides*. Cambridge, UK: Cambridge University Press. pp. 48–60. ISBN 0-521-46827-2.
- [62] Gutteridge A, Thornton JM (November 2005). "Understanding nature's catalytic toolkit". *Trends in Biochemical Sciences* **30** (11): 622–9. doi:10.1016/j.tibs.2005.09.006. PMID 16214343.
- [63] McMurry, John (1996). *Organic chemistry*. Pacific Grove, CA, USA: Brooks/Cole. p. 1064. ISBN 0-534-23832-7.
- [64] Strecker, Adolph (1850). "Ueber die künstliche Bildung der Milchsäure und einen neuen, dem Glycocoll homologen Körper". *Justus Liebigs Annalen der Chemie* **75** (1): 27–45. doi:10.1002/jlac.18500750103.
- [65] Strecker, Adolph (1854). "Ueber einen neuen aus Aldehyd - Ammoniak und Blausäure entstehenden Körper". *Justus Liebigs Annalen der Chemie* **91** (3): 349–51. doi:10.1002/jlac.18540910309.

- [66] Masumoto S, Usuda H, Suzuki M, Kanai M, Shibasaki M (May 2003). "Catalytic enantioselective Strecker reaction of ketoimines". *Journal of the American Chemical Society* **125** (19): 5634–5. doi:10.1021/ja034980. PMID 12733893.
- [67] Davis, F. A. (1994). *Tetrahedron Letters* **35**: 9351.
- [68] Ishitani, Haruro; Komiyama, Susumu; Hasegawa, Yoshiki; Kobayashi, Shū (2000). "Catalytic Asymmetric Strecker Synthesis. Preparation of Enantiomerically Pure α -Amino Acid Derivatives from Aldimines and Tributyltin Cyanide or Achiral Aldehydes, Amines, and Hydrogen Cyanide Using a Chiral Zirconium Catalyst". *Journal of the American Chemical Society* **122** (5): 762–6. doi:10.1021/ja9935207.
- [69] Huang, Jinkun; Corey, E. J. (2004). "A New Chiral Catalyst for the Enantioselective Strecker Synthesis of α -Amino Acids". *Org. Letters* **62** (6): 5027–9. doi:10.1021/ol047698w.
- [70] Duthaler, Rudolf O. (1994). "Recent developments in the stereoselective synthesis of α -amino acids". *Tetrahedron* **50** (6): 1539–1650. doi:10.1016/S0040-4020(01)80840-1.
- [71] Ibba M, Söll D (May 2001). "The renaissance of aminoacyl-tRNA synthesis" (<http://www.nature.com/embor/journal/v2/n5/full/embor420.html>). *EMBO Reports* **2** (5): 382–7. doi:10.1093/embo-reports/kve095 (inactive 2010-02-18). PMID 11375928. PMC 1083889. .
- [72] Lengyel P, Söll D (June 1969). "Mechanism of protein biosynthesis" (<http://mmbr.asm.org/cgi/pmidlookup?view=long&pmid=4896351>). *Bacteriological Reviews* **33** (2): 264–301. PMID 4896351. PMC 378322. .
- [73] Wu G, Fang YZ, Yang S, Lupton JR, Turner ND (March 2004). "Glutathione metabolism and its implications for health" (<http://jn.nutrition.org/cgi/pmidlookup?view=long&pmid=14988435>). *The Journal of Nutrition* **134** (3): 489–92. PMID 14988435. .
- [74] Meister A (November 1988). "Glutathione metabolism and its selective modification" (<http://www.jbc.org/cgi/pmidlookup?view=long&pmid=3053703>). *The Journal of Biological Chemistry* **263** (33): 17205–8. PMID 3053703. .
- [75] Carpino, Louis A. (1992). "1-Hydroxy-7-azabenzotriazole. An efficient peptide coupling additive". *Journal of the American Chemical Society* **115** (10): 4397–8. doi:10.1021/ja00063a082.
- [76] Marasco D, Perretta G, Sabatella M, Ruvo M (October 2008). "Past and future perspectives of synthetic peptide libraries". *Current Protein & Peptide Science* **9** (5): 447–67. doi:10.2174/138920308785915209. PMID 18855697.
- [77] Jones, Russell Celyn; Buchanan, Bob B.; Gruissem, Wilhelm (2000). *Biochemistry & molecular biology of plants*. Rockville, Md: American Society of Plant Physiologists. pp. 371–2. ISBN 0-943088-39-9.
- [78] Stryer, Lubert; Berg, Jeremy Mark; Tymoczko, John L. (2002). *Biochemistry*. San Francisco: W.H. Freeman. pp. 639–49. ISBN 0-7167-4684-0.
- [79] Kivirikko KI, Pihlajaniemi T (1998). "Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases". *Advances in Enzymology and Related Areas of Molecular Biology* **72**: 325–98. PMID 9559057.
- [80] Whitmore L, Wallace BA (May 2004). "Analysis of peptaibol sequence composition: implications for in vivo synthesis and channel formation". *European Biophysics Journal* **33** (3): 233–7. doi:10.1007/s00249-003-0348-1. PMID 14534753.
- [81] Alexander L, Grierson D (October 2002). "Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening". *Journal of Experimental Botany* **53** (377): 2039–55. doi:10.1093/jxb/erf072. PMID 12324528.
- [82] Urry, Dan W. (2004). "The change in Gibbs free energy for hydrophobic association: Derivation and evaluation by means of inverse temperature transitions". *Chemical Physics Letters* **399** (1–3): 177–83. doi:10.1016/S0009-2614(04)01565-9.
- [83] Magee T, Seabra MC (April 2005). "Fatty acylation and prenylation of proteins: what's hot in fat". *Current Opinion in Cell Biology* **17** (2): 190–6. doi:10.1016/j.ceb.2005.02.003. PMID 15780596.
- [84] Pilobello KT, Mahal LK (June 2007). "Deciphering the glycode: the complexity and analytical challenge of glycomics". *Current Opinion in Chemical Biology* **11** (3): 300–5. doi:10.1016/j.cbpa.2007.05.002. PMID 17500024.
- [85] Smotrys JE, Linder ME (2004). "Palmitoylation of intracellular signaling proteins: regulation and function". *Annual Review of Biochemistry* **73**: 559–87. doi:10.1146/annurev.biochem.73.011303.073954. PMID 15189153.
- [86] Hausman, Robert E.; Cooper, Geoffrey M. (2004). *The cell: a molecular approach*. Washington, D.C: ASM Press. p. 51. ISBN 0-87893-214-3.
- [87] Kyte J, Doolittle RF (May 1982). "A simple method for displaying the hydropathic character of a protein". *Journal of Molecular Biology* **157** (1): 105–32. doi:10.1016/0022-2836(82)90515-0. PMID 7108955.
- [88] Freifelder, D. (1983). *Physical Biochemistry* (2nd ed.). W. H. Freeman and Company. ISBN 0-7167-1315-2.
- [89] Suchanek M, Radzikowska A, Thiele C (April 2005). "Photo-leucine and photo-methionine allow identification of protein-protein interactions in living cells". *Nature Methods* **2** (4): 261–7. doi:10.1038/nmeth752. PMID 15782218.
- [90] <http://www.peptideguide.com/amino-acids/index.html>
- [91] <http://www.unc.edu/~bzafer/aminoacids/>
- [92] <http://www.chem.qmul.ac.uk/iupac/AminoAcid/>
- [93] <http://micro.magnet.fsu.edu/aminoacids/index.html>
- [94] <http://www.russell.embl.de/aas/>
- [95] <http://www.organic-chemistry.org/synthesis/C1C/nitrogen/alpha-amino-acids2.shtm>
- [96] <http://www.mathiasbader.de/studium/biology/index.php?lng=en>
- [97] http://www.biology.arizona.edu/biochemistry/problem_sets/aa/Dayhoff.html
- [98] http://www2.iq.usp.br/docente/gutz/Curtipot_.html

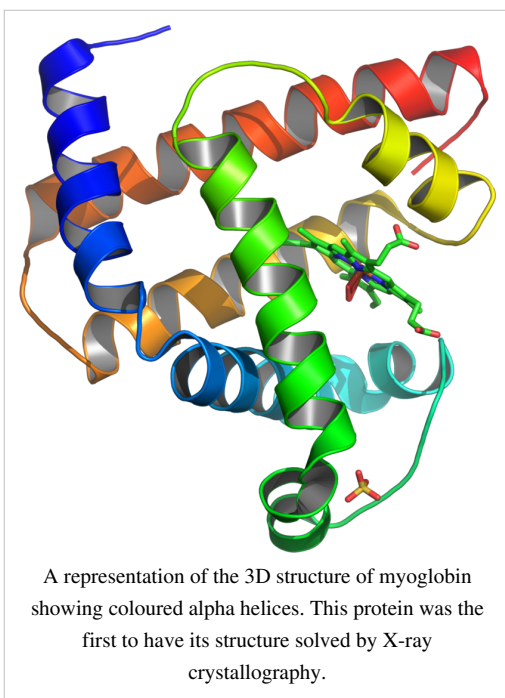
Proteins



This

article is part of the series on:
Gene expression
a Molecular biology topic (*portal*)
(Glossary)

Introduction to Genetics
General flow: DNA > RNA > Protein
special transfers (RNA > RNA, RNA > DNA, Protein > Protein)
Genetic code
Transcription
Transcription (<i>Transcription factors</i> , <i>RNA Polymerase,promoter</i>) Prokaryotic / Archaeal / Eukaryotic
post-transcriptional modification (<i>hnRNA,Splicing</i>)
Translation
Translation (<i>Ribosome,tRNA</i>) Prokaryotic / Archaeal / Eukaryotic
post-translational modification (<i>functional groups, peptides</i> , <i>structural changes</i>)
gene regulation
epigenetic regulation (<i>Genomic imprinting</i>)
transcriptional regulation
post-transcriptional regulation (<i>sequestration</i> , <i>alternative splicing,miRNA</i>)
translational regulation
post-translational regulation (<i>reversible,irreversible</i>)
ask a question ^[1] , edit ^[2]



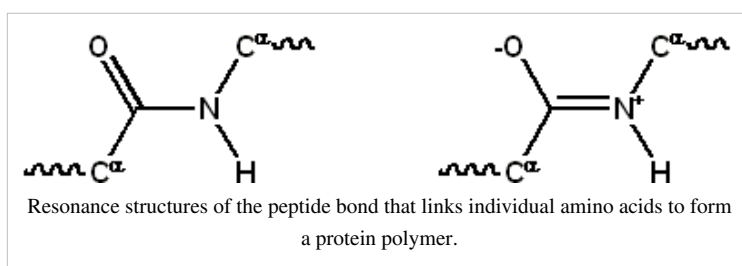
Proteins (also known as **polypeptides**) are organic compounds made of amino acids arranged in a linear chain and folded into a globular form. The amino acids in a polymer are joined together by the peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. The sequence of amino acids in a protein is defined by the sequence of a gene, which is encoded in the genetic code.^[3] In general, the genetic code specifies 20 standard amino acids; however, in certain organisms the genetic code can include selenocysteine—and in certain archaea—pyrrolysine. Shortly after or even during synthesis, the residues in a protein are often chemically modified by post-translational modification, which alters the physical and chemical properties, folding, stability, activity, and ultimately, the function of the proteins. Proteins can also work together to achieve a particular function, and they often associate to form stable complexes.^[4]

Like other biological macromolecules such as polysaccharides and nucleic acids, proteins are essential parts of organisms and participate in virtually every process within cells. Many proteins are enzymes that catalyze biochemical reactions and are vital to metabolism. Proteins also have structural or mechanical functions, such as actin and myosin in muscle and the proteins in the cytoskeleton, which form a system of scaffolding that maintains cell shape. Other proteins are important in cell signaling, immune responses, cell adhesion, and the cell cycle. Proteins are also necessary in animals' diets, since animals cannot synthesize all the amino acids they need and must obtain essential amino acids from food. Through the process of digestion, animals break down ingested protein into free amino acids that are then used in metabolism.

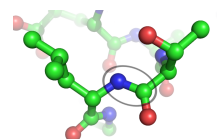
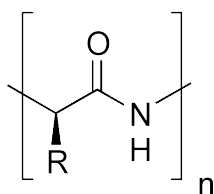
Proteins were first described by the Dutch chemist Gerhardus Johannes Mulder and named by the Swedish chemist Jöns Jakob Berzelius in 1838. The central role of proteins in living organisms was however not fully appreciated until 1926, when James B. Sumner showed that the enzyme urease was a protein.^[5] The first protein to be sequenced was insulin, by Frederick Sanger, who won the Nobel Prize for this achievement in 1958. The first protein structures to be solved were hemoglobin and myoglobin, by Max Perutz and Sir John Cowdery Kendrew, respectively, in 1958.^[6] ^[7] The three-dimensional structures of both proteins were first determined by x-ray diffraction analysis; Perutz and Kendrew shared the 1962 Nobel Prize in Chemistry for these discoveries. Proteins may be purified from other cellular components using a variety of techniques such as ultracentrifugation, precipitation, electrophoresis, and chromatography; the advent of genetic engineering has made possible a number of methods to facilitate purification. Methods commonly used to study protein structure and function include immunohistochemistry, site-directed mutagenesis, and mass spectrometry.

Biochemistry

Most proteins are linear polymers built from series of up to 20 different L- α -amino acids. All amino acids possess common structural features, including an α -carbon to which an amino group, a carboxyl group, and a



variable side chain are bonded. Only proline differs from this basic structure as it contains an unusual ring to the N-end amine group, which forces the CO–NH amide moiety into a fixed conformation.^[8] The side chains of the standard amino acids, detailed in the list of standard amino acids, have a great variety of chemical structures and properties; it is the combined effect of all of the amino acid side chains in a protein that ultimately determines its three-dimensional structure and its chemical reactivity.^[9]

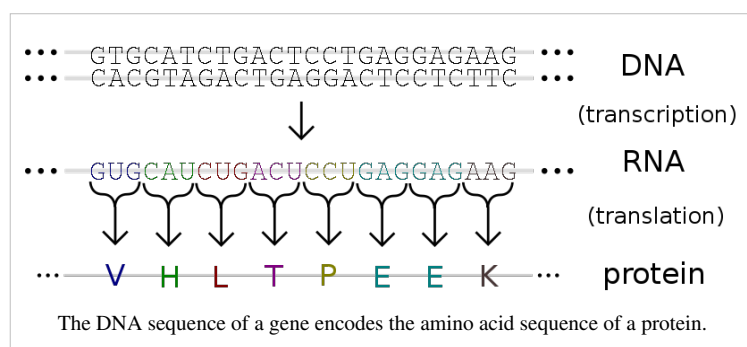


Chemical structure of the peptide bond (left) and a peptide bond between leucine and threonine (right).

The amino acids in a polypeptide chain are linked by peptide bonds. Once linked in the protein chain, an individual amino acid is called a *residue*, and the linked series of carbon, nitrogen, and oxygen atoms are known as the *main chain* or *protein backbone*.^[10] The peptide bond has two resonance forms that contribute some double-bond character and inhibit rotation around its axis, so that the carbons are roughly coplanar. The other two dihedral angles in the peptide bond determine the local shape assumed by the protein backbone.^[11] The end of the protein with a free carboxyl group is known as the C-terminus or carboxy terminus, whereas the end with a free amino group is known as the N-terminus or amino terminus.

The words *protein*, *polypeptide*, and *peptide* are a little ambiguous and can overlap in meaning. *Protein* is generally used to refer to the complete biological molecule in a stable conformation, whereas *peptide* is generally reserved for a short amino acid oligomers often lacking a stable three-dimensional structure. However, the boundary between the two is not well defined and usually lies near 20–30 residues.^[12] *Polypeptide* can refer to any single linear chain of amino acids, usually regardless of length, but often implies an absence of a defined conformation.

Synthesis



Proteins are assembled from amino acids using information encoded in genes. Each protein has its own unique amino acid sequence that is specified by the nucleotide sequence of the gene encoding this protein. The genetic code is a set of three-nucleotide sets called codons and each three-nucleotide combination designates an amino acid, for example AUG (adenine-uracil-guanine) is the code for methionine. Because DNA

contains four nucleotides, the total number of possible codons is 64; hence, there is some redundancy in the genetic code, with some amino acids specified by more than one codon.^[13] Genes encoded in DNA are first transcribed into pre-messenger RNA (mRNA) by proteins such as RNA polymerase. Most organisms then process the pre-mRNA (also known as a *primary transcript*) using various forms of post-transcriptional modification to form the mature mRNA, which is then used as a template for protein synthesis by the ribosome. In prokaryotes the mRNA may either be used as soon as it is produced, or be bound by a ribosome after having moved away from the nucleoid. In contrast, eukaryotes make mRNA in the cell nucleus and then translocate it across the nuclear membrane into the cytoplasm, where protein synthesis then takes place. The rate of protein synthesis is higher in prokaryotes than eukaryotes and can reach up to 20 amino acids per second.^[14]

The process of synthesizing a protein from an mRNA template is known as translation. The mRNA is loaded onto the ribosome and is read three nucleotides at a time by matching each codon to its base pairing anticodon located on a transfer RNA molecule, which carries the amino acid corresponding to the codon it recognizes. The enzyme aminoacyl tRNA synthetase "charges" the tRNA molecules with the correct amino acids. The growing polypeptide is often termed the *nascent chain*. Proteins are always biosynthesized from N-terminus to C-terminus.^[13]

The size of a synthesized protein can be measured by the number of amino acids it contains and by its total molecular mass, which is normally reported in units of *daltons* (synonymous with atomic mass units), or the derivative unit kilodalton (kDa). Yeast proteins are on average 466 amino acids long and 53 kDa in mass.^[12] The largest known proteins are the titins, a component of the muscle sarcomere, with a molecular mass of almost 3,000 kDa and a total length of almost 27,000 amino acids.^[15]

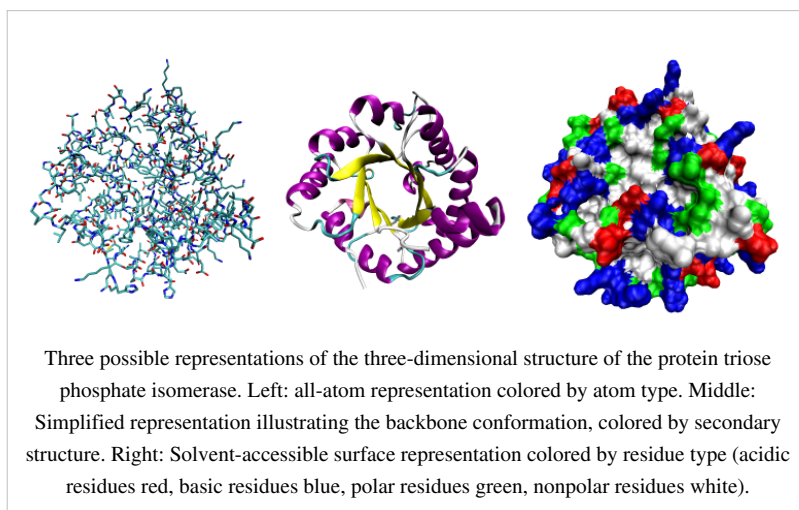
Chemical synthesis

Short proteins can also be synthesized chemically by a family of methods known as peptide synthesis, which rely on organic synthesis techniques such as chemical ligation to produce peptides in high yield.^[16] Chemical synthesis allows for the introduction of non-natural amino acids into polypeptide chains, such as attachment of fluorescent probes to amino acid side chains.^[17] These methods are useful in laboratory biochemistry and cell biology, though generally not for commercial applications. Chemical synthesis is inefficient for polypeptides longer than about 300 amino acids, and the synthesized proteins may not readily assume their native tertiary structure. Most chemical synthesis methods proceed from C-terminus to N-terminus, opposite the biological reaction.^[18]

Structure of proteins

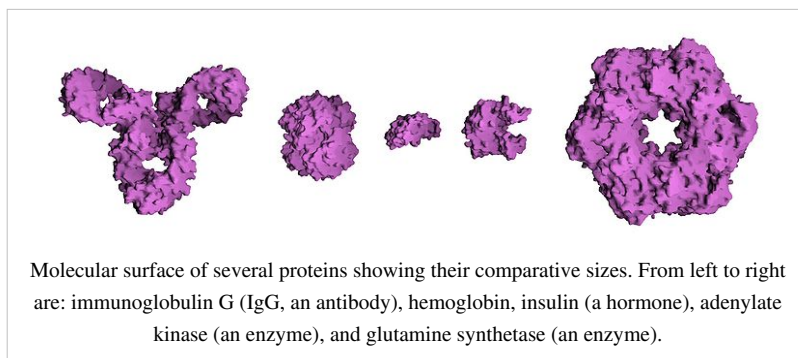
Most proteins fold into unique 3-dimensional structures. The shape into which a protein naturally folds is known as its native conformation.^[19] Although many proteins can fold unassisted, simply through the chemical properties of their amino acids, others require the aid of molecular chaperones to fold into their native states.^[20] Biochemists often refer to four distinct aspects of a protein's structure:^[21]

- **Primary structure:** the amino acid sequence.
- **Secondary structure:** regularly repeating local structures stabilized by hydrogen bonds. The most common examples are the alpha helix, beta sheet and turns. Because secondary structures are local, many regions of different secondary structure can be present in the same protein molecule.
- **Tertiary structure:** the overall shape of a single protein molecule; the spatial relationship of the secondary structures to one another. Tertiary structure is generally stabilized by nonlocal interactions, most commonly the formation of a hydrophobic core, but also through salt bridges, hydrogen bonds, disulfide bonds, and even post-translational modifications. The term "tertiary structure" is often used as synonymous with the term *fold*. The Tertiary structure is what controls the basic function of the protein.
- **Quaternary structure:** the structure formed by several protein molecules (polypeptide chains), usually called *protein subunits* in this context, which function as a single protein complex.



Proteins are not entirely rigid molecules. In addition to these levels of structure, proteins may shift between several related structures while they perform their functions. In the context of these functional rearrangements, these tertiary or quaternary structures are usually referred to as "conformations", and transitions between them are called *conformational changes*. Such changes are often induced by the binding of a substrate molecule to an enzyme's active site, or the physical region of the protein that participates in chemical catalysis. In solution proteins also undergo variation in structure through thermal vibration and the collision with other molecules.^[22]

Proteins can be informally divided into three main classes, which correlate with typical tertiary structures: globular proteins, fibrous proteins, and membrane proteins. Almost all globular proteins are soluble and many are enzymes. Fibrous proteins are often structural, such as collagen, the major component of connective tissue, or keratin, the protein component of hair and nails. Membrane proteins often serve as receptors or provide channels for polar or charged molecules to pass through the cell membrane.^[23]



A special case of intramolecular hydrogen bonds within proteins, poorly shielded from water attack and hence promoting their own dehydration, are called dehydrons.^[24]

Structure determination

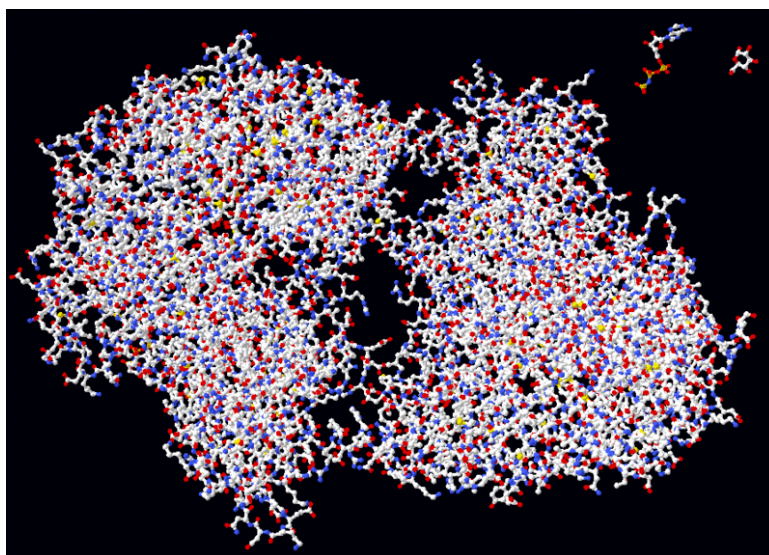
Discovering the tertiary structure of a protein, or the quaternary structure of its complexes, can provide important clues about how the protein performs its function. Common experimental methods of structure determination include X-ray crystallography and NMR spectroscopy, both of which can produce information at atomic resolution. Dual polarisation interferometry is a quantitative analytical method for measuring the overall protein conformation and conformational changes due to interactions or other stimulus. Circular dichroism is another laboratory technique for determining internal beta sheet/ helical composition of proteins. Cryoelectron microscopy is used to produce lower-resolution structural information about very large protein complexes, including assembled viruses;^[25] a variant known as electron crystallography can also produce high-resolution information in some cases, especially for two-dimensional crystals of membrane proteins.^[26] Solved structures are usually deposited in the Protein Data Bank (PDB), a freely available resource from which structural data about thousands of proteins can be obtained in the form of Cartesian coordinates for each atom in the protein.^[27]

Many more gene sequences are known than protein structures. Further, the set of solved structures is biased toward proteins that can be easily subjected to the conditions required in X-ray crystallography, one of the major structure determination methods. In particular, globular proteins are comparatively easy to crystallize in preparation for X-ray crystallography. Membrane proteins, by contrast, are difficult to crystallize and are underrepresented in the PDB.^[28] Structural genomics initiatives have attempted to remedy these deficiencies by systematically solving representative structures of major fold classes. Protein structure prediction methods attempt to provide a means of generating a plausible structure for proteins whose structures have not been experimentally determined.

Cellular functions

Proteins are the chief actors within the cell, said to be carrying out the duties specified by the information encoded in genes.^[12] With the exception of certain types of RNA, most other biological molecules are relatively inert elements upon which proteins act. Proteins make up half the dry weight of an *Escherichia coli* cell, whereas other macromolecules such as DNA and RNA make up only 3% and 20%, respectively.^[29] The set of proteins expressed in a particular cell or cell type is known as its proteome.

The chief characteristic of proteins that also allows their diverse set of functions is their ability to bind other molecules specifically and tightly. The region of the protein responsible for binding another molecule is known as the binding site and is often a depression or "pocket" on the molecular surface. This binding ability is mediated by the tertiary structure of the protein, which defines the binding site pocket, and by the chemical properties of the surrounding amino acids' side chains. Protein binding can be extraordinarily tight and specific; for example, the ribonuclease inhibitor protein binds to human angiogenin



The enzyme hexokinase is shown as a simple ball-and-stick molecular model. To scale in the top right-hand corner are two of its substrates, ATP and glucose.

with a sub-femtomolar dissociation constant ($<10^{-15}$ M) but does not bind at all to its amphibian homolog onconase (>1 M). Extremely minor chemical changes such as the addition of a single methyl group to a binding partner can sometimes suffice to nearly eliminate binding; for example, the aminoacyl tRNA synthetase specific to the amino acid valine discriminates against the very similar side chain of the amino acid isoleucine.^[30]

Proteins can bind to other proteins as well as to small-molecule substrates. When proteins bind specifically to other copies of the same molecule, they can oligomerize to form fibrils; this process occurs often in structural proteins that consist of globular monomers that self-associate to form rigid fibers. Protein–protein interactions also regulate enzymatic activity, control progression through the cell cycle, and allow the assembly of large protein complexes that carry out many closely related reactions with a common biological function. Proteins can also bind to, or even be integrated into, cell membranes. The ability of binding partners to induce conformational changes in proteins allows the construction of enormously complex signaling networks.^[31] Importantly, as interactions between proteins are reversible, and depend heavily on the availability of different groups of partner proteins to form aggregates that are capable to carry out discrete sets of function, study of the interactions between specific proteins is a key to understand important aspects of cellular function, and ultimately the properties that distinguish particular cell types^{[32] [33]}.

Enzymes

The best-known role of proteins in the cell is as enzymes, which catalyze chemical reactions. Enzymes are usually highly specific and accelerate only one or a few chemical reactions. Enzymes carry out most of the reactions involved in metabolism, as well as manipulating DNA in processes such as DNA replication, DNA repair, and transcription. Some enzymes act on other proteins to add or remove chemical groups in a process known as post-translational modification. About 4,000 reactions are known to be catalyzed by enzymes.^[34] The rate acceleration conferred by enzymatic catalysis is often enormous — as much as 10^{17} -fold increase in rate over the

uncatalyzed reaction in the case of orotate decarboxylase (78 million years without the enzyme, 18 milliseconds with the enzyme).^[35]

The molecules bound and acted upon by enzymes are called substrates. Although enzymes can consist of hundreds of amino acids, it is usually only a small fraction of the residues that come in contact with the substrate, and an even smaller fraction — 3 to 4 residues on average — that are directly involved in catalysis.^[36] The region of the enzyme that binds the substrate and contains the catalytic residues is known as the active site.

Cell signaling and ligand binding

Many proteins are involved in the process of cell signaling and signal transduction. Some proteins, such as insulin, are extracellular proteins that transmit a signal from the cell in which they were synthesized to other cells in distant tissues. Others are membrane proteins that act as receptors whose main function is to bind a signaling molecule and induce a biochemical response in the cell. Many receptors have a binding site exposed on the cell surface and an effector domain within the cell, which may have enzymatic activity or may undergo a conformational change detected by other proteins within the cell.^[37]

Antibodies are protein components of adaptive immune system whose main function is to bind antigens, or foreign substances in the body, and target them for destruction. Antibodies can be secreted into the extracellular environment or anchored in the membranes of specialized B cells known as plasma cells. Whereas enzymes are limited in their binding affinity for their substrates by the necessity of conducting their reaction, antibodies have no such constraints. An antibody's binding affinity to its target is extraordinarily high.^[38]

Many ligand transport proteins bind particular small biomolecules and transport them to other locations in the body of a multicellular organism. These proteins must have a high binding affinity when their ligand is present in high concentrations, but must also release the ligand when it is present at low concentrations in the target tissues. The canonical example of a ligand-binding protein is haemoglobin, which transports oxygen from the lungs to other organs and tissues in all vertebrates and has close homologs in every biological kingdom.^[39] Lectins are sugar-binding proteins which are highly specific for their sugar moieties. Lectins typically play a role in biological recognition phenomena involving cells and proteins.^[40] Receptors and hormones are highly specific binding proteins.

Transmembrane proteins can also serve as ligand transport proteins that alter the permeability of the cell membrane to small molecules and ions. The membrane alone has a hydrophobic core through which polar or charged molecules cannot diffuse. Membrane proteins contain internal channels that allow such molecules to enter and exit the cell. Many ion channel proteins are specialized to select for only a particular ion; for example, potassium and sodium channels often discriminate for only one of the two ions.^[41]

Structural proteins

Structural proteins confer stiffness and rigidity to otherwise-fluid biological components. Most structural proteins are fibrous proteins; for example, actin and tubulin are globular and soluble as monomers, but polymerize to form long, stiff fibers that comprise the cytoskeleton, which allows the cell to maintain its shape and size. Collagen and elastin are critical components of connective tissue such as cartilage, and keratin is found in hard or filamentous structures such as hair, nails, feathers, hooves, and some animal shells.^[42]



Other proteins that serve structural functions are motor proteins such as myosin, kinesin, and dynein, which are capable of generating mechanical forces. These proteins are crucial for cellular motility of single celled organisms and the sperm of many multicellular organisms which reproduce sexually. They also generate the forces exerted by contracting muscles.^[43]

Methods of study

As some of the most commonly studied biological molecules, the activities and structures of proteins are examined both *in vitro* and *in vivo*. *In vitro* studies of purified proteins in controlled environments are useful for learning how a protein carries out its function: for example, enzyme kinetics studies explore the chemical mechanism of an enzyme's catalytic activity and its relative affinity for various possible substrate molecules. By contrast, *in vivo* experiments on proteins' activities within cells or even within whole organisms can provide complementary information about where a protein functions and how it is regulated.

Protein purification

In order to perform *in vitro* analysis, a protein must be purified away from other cellular components. This process usually begins with cell lysis, in which a cell's membrane is disrupted and its internal contents released into a solution known as a crude lysate. The resulting mixture can be purified using ultracentrifugation, which fractionates the various cellular components into fractions containing soluble proteins; membrane lipids and proteins; cellular organelles, and nucleic acids. Precipitation by a method known as salting out can concentrate the proteins from this lysate. Various types of chromatography are then used to isolate the protein or proteins of interest based on properties such as molecular weight, net charge and binding affinity.^[44] The level of purification can be monitored using various types of gel electrophoresis if the desired protein's molecular weight and isoelectric point are known, by spectroscopy if the protein has distinguishable spectroscopic features, or by enzyme assays if the protein has enzymatic activity. Additionally, proteins can be isolated according their charge using electrofocusing.^[45]

For natural proteins, a series of purification steps may be necessary to obtain protein sufficiently pure for laboratory applications. To simplify this process, genetic engineering is often used to add chemical features to proteins that make them easier to purify without affecting their structure or activity. Here, a "tag" consisting of a specific amino acid sequence, often a series of histidine residues (a "His-tag"), is attached to one terminus of the protein. As a result, when the lysate is passed over a chromatography column containing nickel, the histidine residues ligate the nickel and attach to the column while the untagged components of the lysate pass unimpeded. A number of different tags have been developed to help researchers purify specific proteins from complex mixtures.^[46]

Cellular localization

The study of proteins *in vivo* is often concerned with the synthesis and localization of the protein within the cell. Although many intracellular proteins are synthesized in the cytoplasm and membrane-bound or secreted proteins in the endoplasmic reticulum, the specifics of how proteins are targeted to specific organelles or cellular structures is often unclear. A useful technique for assessing cellular localization uses genetic engineering to express in a cell a fusion protein or chimera consisting of the natural protein of interest linked to a "reporter" such as green fluorescent protein (GFP).^[47] The fused protein's position within the cell can be cleanly and efficiently visualized using microscopy,^[48] as shown in the figure opposite.

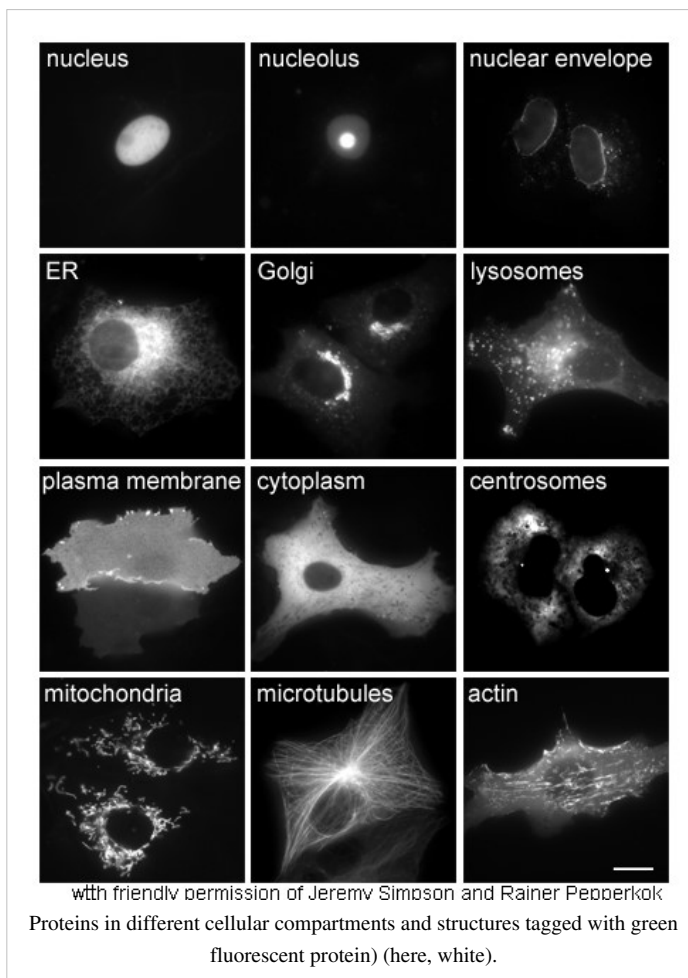
Other methods for elucidating the cellular location of proteins requires the use of known compartmental markers for regions such as the ER, the Golgi, lysosomes/vacuoles, mitochondria, chloroplasts, plasma membrane, etc. With the use of fluorescently-tagged versions of these markers or of antibodies to known markers, it becomes much simpler to identify the localization of a protein of interest.

For example, indirect immunofluorescence will allow for fluorescence colocalization and demonstration of location. Fluorescent dyes are used to label cellular compartments for a similar purpose.^[49]

Other possibilities exist, as well. For example, immunohistochemistry usually utilizes an antibody to one or more proteins of interest that are conjugated to enzymes yielding either luminescent or chromogenic signals that can be compared between samples, allowing for localization information. Another applicable technique is cofractionation in sucrose (or other material) gradients using isopycnic centrifugation.^[50] While this technique does not prove colocalization of a compartment of known density and the protein of interest, it does increase the likelihood, and is more amenable to large-scale studies.

Finally, the gold-standard method of cellular localization is immunoelectron microscopy. This technique also uses an antibody to the protein of interest, along with classical electron microscopy techniques. The sample is prepared for normal electron microscopic examination, and then treated with an antibody to the protein of interest that is conjugated to an extremely electro-dense material, usually gold. This allows for the localization of both ultrastructural details as well as the protein of interest.^[51]

Through another genetic engineering application known as site-directed mutagenesis, researchers can alter the protein sequence and hence its structure, cellular localization, and susceptibility to regulation. This technique even allows the incorporation of unnatural amino acids into proteins, using modified tRNAs,^[52] and may allow the rational design of new proteins with novel properties.^[53]



Proteomics and bioinformatics

The total complement of proteins present at a time in a cell or cell type is known as its proteome, and the study of such large-scale data sets defines the field of proteomics, named by analogy to the related field of genomics. Key experimental techniques in proteomics include 2D electrophoresis,^[54] which allows the separation of a large number of proteins, mass spectrometry,^[55] which allows rapid high-throughput identification of proteins and sequencing of peptides (most often after in-gel digestion), protein microarrays,^[56] which allow the detection of the relative levels of a large number of proteins present in a cell, and two-hybrid screening, which allows the systematic exploration of protein–protein interactions.^[57] The total complement of biologically possible such interactions is known as the interactome.^[58] A systematic attempt to determine the structures of proteins representing every possible fold is known as structural genomics.^[59]

The large amount of genomic and proteomic data available for a variety of organisms, including the human genome, allows researchers to efficiently identify homologous proteins in distantly related organisms by sequence alignment. Sequence profiling tools can perform more specific sequence manipulations such as restriction enzyme maps, open reading frame analyses for nucleotide sequences, and secondary structure prediction. From this data phylogenetic trees can be constructed and evolutionary hypotheses developed using special software like ClustalW regarding the ancestry of modern organisms and the genes they express. The field of bioinformatics seeks to assemble, annotate, and analyze genomic and proteomic data, applying computational techniques to biological problems such as gene finding and cladistics.

Structure prediction and simulation

Complementary to the field of structural genomics, protein structure prediction seeks to develop efficient ways to provide plausible models for proteins whose structures have not yet been determined experimentally^[60]. The most successful type of structure prediction, known as homology modeling, relies on the existence of a "template" structure with sequence similarity to the protein being modeled; structural genomics' goal is to provide sufficient representation in solved structures to model most of those that remain.^[61] Although producing accurate models remains a challenge when only distantly related template structures are available, it has been suggested that sequence alignment is the bottleneck in this process, as quite accurate models can be produced if a "perfect" sequence alignment is known.^[62] Many structure prediction methods have served to inform the emerging field of protein engineering, in which novel protein folds have already been designed.^[63] A more complex computational problem is the prediction of intermolecular interactions, such as in molecular docking and protein–protein interaction prediction.^[64]

The processes of protein folding and binding can be simulated using such technique as molecular mechanics, in particular, molecular dynamics and Monte Carlo, which increasingly take advantage of parallel and distributed computing (Folding@Home project^[65]; molecular modeling on GPU). The folding of small alpha-helical protein domains such as the villin headpiece^[66] and the HIV accessory protein^[67] have been successfully simulated *in silico*, and hybrid methods that combine standard molecular dynamics with quantum mechanics calculations have allowed exploration of the electronic states of rhodopsins.^[68]

Nutrition

Most microorganisms and plants can biosynthesize all 20 standard amino acids, while animals (including humans) must obtain some of the amino acids from the diet.^[29] The amino acids that an organism cannot synthesize on its own are referred to as essential amino acids. Key enzymes that synthesize certain amino acids are not present in animals — such as aspartokinase, which catalyzes the first step in the synthesis of lysine, methionine, and threonine from aspartate. If amino acids are present in the environment, microorganisms can conserve energy by taking up the amino acids from their surroundings and downregulating their biosynthetic pathways.

In animals, amino acids are obtained through the consumption of foods containing protein. Ingested proteins are then broken down into amino acids through digestion, which typically involves denaturation of the protein through exposure to acid and hydrolysis by enzymes called proteases. Some ingested amino acids are used for protein biosynthesis, while others are converted to glucose through gluconeogenesis, or fed into the citric acid cycle. This use of protein as a fuel is particularly important under starvation conditions as it allows the body's own proteins to be used to support life, particularly those found in muscle.^[69] Amino acids are also an important dietary source of nitrogen.

History and etymology

Proteins were recognized as a distinct class of biological molecules in the eighteenth century by Antoine Fourcroy and others, distinguished by the molecules' ability to coagulate or flocculate under treatments with heat or acid. Noted examples at the time included albumin from egg whites, blood serum albumin, fibrin, and wheat gluten. Dutch chemist Gerhardus Johannes Mulder carried out elemental analysis of common proteins and found that nearly all proteins had the same empirical formula, $C_{400}H_{620}N_{100}O_{120}P_1S_1$.^[70] He came to the erroneous conclusion that they might be composed of a single type of (very large) molecule. The term "protein" to describe these molecules was proposed in 1838 by Mulder's associate Jöns Jakob Berzelius; protein is derived from the Greek word *πρωτεῖος* (*proteios*), meaning "primary"^[71], "in the lead", or "standing in front".^[72] Mulder went on to identify the products of protein degradation such as the amino acid leucine for which he found a (nearly correct) molecular weight of 131 Da.^[70]

The difficulty in purifying proteins in large quantities made them very difficult for early protein biochemists to study. Hence, early studies focused on proteins that could be purified in large quantities, e.g., those of blood, egg white, various toxins, and digestive/metabolic enzymes obtained from slaughterhouses. In the 1950s, the Armour Hot Dog Co. purified 1 kg of pure bovine pancreatic ribonuclease A and made it freely available to scientists; this gesture helped ribonuclease A become a major target for biochemical study for the following decades.^[70]

Linus Pauling is credited with the successful prediction of regular protein secondary structures based on hydrogen bonding, an idea first put forth by William Astbury in 1933.^[73] Later work by Walter Kauzmann on denaturation,^[74]^[75] based partly on previous studies by Kaj Linderstrøm-Lang,^[76] contributed an understanding of protein folding and structure mediated by hydrophobic interactions. In 1949 Fred Sanger correctly determined the amino acid sequence of insulin, thus conclusively demonstrating that proteins consisted of linear polymers of amino acids rather than branched chains, colloids, or cyclols.^[77] The first atomic-resolution structures of proteins were solved by X-ray crystallography in the 1960s and by NMR in the 1980s. As of 2009, the Protein Data Bank has over 55,000 atomic-resolution structures of proteins.^[78] In more recent times, cryo-electron microscopy of large macromolecular assemblies^[79] and computational protein structure prediction of small protein domains^[80] are two methods approaching atomic resolution.

See also

- Expression cloning
- Intein
- List of proteins
- List of recombinant proteins
- Prion
- Protein design
- Protein dynamics
- Protein structure prediction software
- Proteopathy
- Proteopedia

- Cdx protein family

References

- Branden C, Tooze J. (1999). *Introduction to Protein Structure*. New York: Garland Pub. ISBN 0-8153-2305-0.
- Murray RF, Harper HW, Granner DK, Mayes PA, Rodwell VW. (2006). *Harper's Illustrated Biochemistry*. New York: Lange Medical Books/McGraw-Hill. ISBN 0-07-146197-3.
- Van Holde KE, Mathews CK. (1996). *Biochemistry*. Menlo Park, Calif: Benjamin/Cummings Pub. Co., Inc. ISBN 0-8053-3931-0.
- Jörg von Hagen, VCH-Wiley 2008 Proteomics Sample Preparation. ISBN 978-3-527-31796-7

External links

- Protein Songs (Stuart Mitchell – DNA Music Project) ^[81], 'When a "tape" of mRNA passes through the "playing head" of a ribosome, the "notes" produced are amino acids and the pieces of music they make up are proteins.'

Databases and projects

- Comparative Toxicogenomics Database ^[82] curates protein–chemical interactions, as well as gene/protein–disease relationships and chemical–disease relationships.
- Bioinformatic Harvester ^[83] A Meta search engine (29 databases) for gene and protein information.
- The Protein Databank ^[84] (see also PDB Molecule of the Month ^[85], presenting short accounts on selected proteins from the PDB)
- Proteopedia – Life in 3D ^[86]: rotatable, zoomable 3D model with wiki annotations for every known protein molecular structure.
- UniProt the Universal Protein Resource ^[87]
- The Protein Naming Utility ^[88]
- Human Protein Atlas ^[89]
- NCBI Entrez Protein database ^[90]
- NCBI Protein Structure database ^[91]
- Human Protein Reference Database ^[92]
- Human Proteinpedia ^[93]
- Folding@Home (Stanford University) ^[94]

Tutorials and educational websites

- "An Introduction to Proteins" ^[95] from HOPES (Huntington's Disease Outreach Project for Education at Stanford)
- Proteins: Biogenesis to Degradation – The Virtual Library of Biochemistry and Cell Biology ^[96]

References

- [1] http://en.wikipedia.org/w/index.php?title=Wikipedia_talk:WikiProject_Molecular_and_Cellular_Biology&action=edit§ion=new
- [2] <http://en.wikipedia.org/w/index.php?title=Template:MolBioGeneExp&action=edit>
- [3] Ridley, M. (2006). *Genome*. New York, NY: Harper Perennial. ISBN 0-06-019497-9
- [4] Maton A, Hopkins J, McLaughlin CW, Johnson S, Warner MQ, LaHart D, Wright JD. (1993). *Human Biology and Health*. Englewood Cliffs, Michigan, USA: Prentice Hall. ISBN 0-13-981176-1. OCLC 32308337.
- [5] Sumner, JB. (1926). "The isolation and crystallization of the enzyme urease. Preliminary paper" (<http://www.jbc.org/cgi/reprint/69/2/435.pdf?ijkey=028d5e540dab50accbf86e01be08db51ef49008f>). *Journal of Biological Chemistry* **69**: 435–41. .
- [6] Muirhead H, Perutz M. (1963). "Structure of hemoglobin. A three-dimensional fourier synthesis of reduced human hemoglobin at 5.5 Å resolution". *Nature* **199** (4894): 633–38. doi:10.1038/199633a0. PMID 14074546.
- [7] Kendrew J, Bodo G, Dintzis H, Parrish R, Wyckoff H, Phillips D. (1958). "A three-dimensional model of the myoglobin molecule obtained by x-ray analysis". *Nature* **181** (4610): 662–66. doi:10.1038/181662a0. PMID 13517261.
- [8] Nelson DL, Cox MM. (2005). *Lehninger's Principles of Biochemistry, 4th Edition*. W. H. Freeman and Company, New York.

- [9] Gutteridge A, Thornton JM. (2005). "Understanding nature's catalytic toolkit". *Trends in Biochemical Sciences* **30** (11): 622–29. doi:10.1016/j.tibs.2005.09.006. PMID 16214343.
- [10] Murray *et al.*, p. 19.
- [11] Murray *et al.*, p. 31.
- [12] Lodish H, Berk A, Matsudaira P, Kaiser CA, Krieger M, Scott MP, Zipurksy SL, Darnell J. (2004). *Molecular Cell Biology* 5th ed. WH Freeman and Company: New York, NY.
- [13] van Holde and Mathews, pp. 1002–42.
- [14] Dobson CM. (2000). "The nature and significance of protein folding". in Pain RH. (ed.). *Mechanisms of Protein Folding*. Oxford, Oxfordshire: Oxford University Press. ISBN 0-19-963789-X.
- [15] Fulton A, Isaacs W. (1991). "Titin, a huge, elastic sarcomeric protein with a probable role in morphogenesis". *Bioessays* **13** (4): 157–61. doi:10.1002/bies.950130403. PMID 1859393.
- [16] Bruckdorfer T, Marder O, Albericio F. (2004). "From production of peptides in milligram amounts for research to multi-tons quantities for drugs of the future". *Current Pharmaceutical Biotechnology* **5** (1): 29–43. doi:10.2174/1389201043489620. PMID 14965208.
- [17] Schwarzer D, Cole P. (2005). "Protein semisynthesis and expressed protein ligation: chasing a protein's tail". *Current Opinions in Chemical Biology* **9** (6): 561–69. doi:10.1016/j.cbpa.2005.09.018. PMID 16226484.
- [18] Kent SB. (2009). "Total chemical synthesis of proteins". *Chemical Society Reviews* **38** (2): 338–51. doi:10.1039/b700141j. PMID 19169452.
- [19] Murray *et al.*, p. 36.
- [20] Murray *et al.*, p. 37.
- [21] Murray *et al.*, pp. 30–34.
- [22] van Holde and Mathews, pp. 368–75.
- [23] van Holde and Mathews, pp. 165–85.
- [24] Fernández A, Scott R. (2003). "Dehydron: a structurally encoded signal for protein interaction" ([http://linkinghub.elsevier.com/retrieve/pii/S0006-3495\(03\)74619-0](http://linkinghub.elsevier.com/retrieve/pii/S0006-3495(03)74619-0)). *Biophysical Journal* **85** (3): 1914–28. doi:10.1016/S0006-3495(03)74619-0. PMID 12944304. PMC 1303363. .
- [25] Branden and Tooze, pp. 340–41.
- [26] Gonen T, Cheng Y, Sliz P, Hiroaki Y, Fujiyoshi Y, Harrison SC, Walz T. (2005). "Lipid-protein interactions in double-layered two-dimensional AQP0 crystals" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1350984>). *Nature* **438** (7068): 633–38. doi:10.1038/nature04321. PMID 16319884. PMC 1350984.
- [27] Standley DM, Kinjo AR, Kinoshita K, Nakamura H. (2008). "Protein structure databases with new web services for structural biology and biomedical research" (<http://bib.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=18430752>). *Briefings in Bioinformatics* **9** (4): 276–85. doi:10.1093/bib/bbn015. PMID 18430752. .
- [28] Walian P, Cross TA, Jap BK (2004). "Structural genomics of membrane proteins" (<http://genomebiology.com/1465-6906/5/215>). *Genome Biology* **5** (4): 215. doi:10.1186/gb-2004-5-4-215. PMID 15059248. PMC 395774. . Retrieved 2009-04-14.
- [29] Voet D, Voet JG. (2004). *Biochemistry* Vol 1 3rd ed. Wiley: Hoboken, NJ.
- [30] Sankaranarayanan R, Moras D (2001). "The fidelity of the translation of the genetic code". *Acta Biochimica Polonica* **48** (2): 323–35. PMID 11732604.
- [31] van Holde and Mathews, pp. 830–49.
- [32] Copland JA, Sheffield-Moore M, Koldzic-Zivanovic N, Gentry S, Lamprou G, Tzortzatou-Stathopoulou F, Zoumpourlis V, Urban RJ, Vlahopoulos SA. Sex steroid receptors in skeletal differentiation and epithelial neoplasia: is tissue-specific intervention possible? *Bioessays*. 2009 Jun;31(6):629–641.PMID: 19382224
- [33] Samarin S, Nusrat A. Regulation of epithelial apical junctional complex by Rho family GTPases. *Front Biosci*. 2009 Jan 1;14:1129–42. Review. PMID: 19273120
- [34] Bairoch A. (2000). "The ENZYME database in 2000" (<http://www.expasy.org/NAR/enz00.pdf>). *Nucleic Acids Research* **28** (1): 304–305. doi:10.1093/nar/28.1.304. PMID 10592255. PMC 102465. .
- [35] Radzicka A, Wolfenden R. (1995). "A proficient enzyme". *Science* **6** (267): 90–93. doi:10.1126/science.7809611. PMID 7809611.
- [36] The Catalytic Site Atlas at The European Bioinformatics Institute (<http://www.ebi.ac.uk/thornton-srv/databases/CSA/>)
- [37] Branden and Tooze, pp. 251–281.
- [38] van Holde and Mathews, pp. 247–50.
- [39] van Holde and Mathews, pp. 220–29.
- [40] Rüdiger H, Siebert HC, Solís D, Jiménez-Barbero J, Romero A, von der Lieth CW, Diaz-Mariño T, Gabius HJ. (2000). "Medicinal chemistry based on the sugar code: fundamentals of lectinology and experimental strategies with lectins as targets" (<http://www.bentham-direct.org/pages/content.php?CMC/2000/00000007/00000004/0002C.SGM>). *Current Medicinal Chemistry* **7** (4): 389–416. PMID 10702616. . Retrieved 2009-04-14.
- [41] Branden and Tooze, pp. 232–234.
- [42] van Holde and Mathews, pp. 178–81.
- [43] van Holde and Mathews, pp. 258–64; 272.
- [44] Murray *et al.*, pp. 21–24.
- [45] Hey J, Posch A, Cohen A, Liu N, Harbers A. (2008). "Fractionation of complex protein mixtures by liquid-phase isoelectric focusing". *Methods in Molecular Biology* **424**: 225–39. doi:10.1007/978-1-60327-064-9_19. PMID 18369866.

- [46] Terpe K. (2003). "Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems". *Applied Microbiology and Biotechnology* **60** (5): 523–33. doi:10.1007/s00253-002-1158-6. PMID 12536251.
- [47] Stepanenko OV, Verkhusha VV, Kuznetsova IM, Uversky VN, Turoverov KK (August 2008). "Fluorescent proteins as biomarkers and biosensors: throwing color lights on molecular and cellular processes". *Curr. Protein Pept. Sci.* **9** (4): 338–69. doi:10.2174/138920308785132668. PMID 18691124.
- [48] Yuste R. (2005). "Fluorescence microscopy today". *Nature Methods* **2** (12): 902–904. doi:10.1038/nmeth1205-902. PMID 16299474.
- [49] Margolin W. (2000). "Green fluorescent protein as a reporter for macromolecular localization in bacterial cells" ([http://linkinghub.elsevier.com/retrieve/pii/S1046-2023\(99\)90906-4](http://linkinghub.elsevier.com/retrieve/pii/S1046-2023(99)90906-4)). *Methods (San Diego, Calif.)* **20** (1): 62–72. doi:10.1006/meth.1999.0906. PMID 10610805. .
- [50] Walker JH, Wilson K. (2000). *Principles and Techniques of Practical Biochemistry*. Cambridge, UK: Cambridge University Press. pp. 287–89. ISBN 0-521-65873-X.
- [51] Mayhew TM, Lucocq JM. (2008). "Developments in cell biology for quantitative immunoelectron microscopy based on thin sections: a review" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2491712>). *Histochemistry and Cell Biology* **130** (2): 299–313. doi:10.1007/s00418-008-0451-6. PMID 18553098. PMC 2491712.
- [52] Hohsaka T, Sisido M (December 2002). "Incorporation of non-natural amino acids into proteins". *Curr Opin Chem Biol* **6** (6): 809–15. doi:10.1016/S1367-5931(02)00376-9. PMID 12470735.
- [53] Cedrone F, Ménez A, Quéménéur E (August 2000). "Tailoring new enzyme functions by rational redesign". *Curr. Opin. Struct. Biol.* **10** (4): 405–10. doi:10.1016/S0959-440X(00)00106-8. PMID 10981626.
- [54] Görg A, Weiss W, Dunn MJ. (2004). "Current two-dimensional electrophoresis technology for proteomics". *Proteomics* **4** (12): 3665–85. doi:10.1002/pmic.200401031. PMID 15543535.
- [55] Conrotto P, Souchelnytskyi S. (2008). "Proteomic approaches in biological and medical sciences: principles and applications" (<http://www.exp-oncology.com.ua/en/archives/36/699.html>). *Experimental Oncology* **30** (3): 171–80. PMID 18806738. .
- [56] Joos T, Bachmann J. (2009). "Protein microarrays: potentials and limitations" (<http://www.bioscience.org/2009/v14/af/3534/fulltext.htm>). *Frontiers in Bioscience* **14**: 4376–85. doi:10.2741/3534. PMID 19273356. .
- [57] Koegl M, Uetz P. (2007). "Improving yeast two-hybrid screening systems" (<http://bfgp.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=18218650>). *Briefings in Functional Genomics & Proteomics* **6** (4): 302–12. doi:10.1093/bfgp/elm035. PMID 18218650. .
- [58] Plewczynski D, Ginalski K. (2009). "The interactome: predicting the protein–protein interactions in cells". *Cellular & Molecular Biology Letters* **14** (1): 1–22. doi:10.2478/s11658-008-0024-7. PMID 18839074.
- [59] Zhang C, Kim SH. (2003). "Overview of structural genomics: from structure to function" (<http://linkinghub.elsevier.com/retrieve/pii/S1367593102000157>). *Current Opinion in Chemical Biology* **7** (1): 28–32. doi:10.1016/S1367-5931(02)00015-7. PMID 12547423. .
- [60] Zhang Y. (2008). "Progress and challenges in protein structure prediction" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2680823>). *Current Opinions in Structural Biology* **18** (3): 342–48. doi:10.1016/j.sbi.2008.02.004. Entrez Pubmed 18436442 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=18436442). PMID 18436442. PMC 2680823.
- [61] Xiang Z. (2006). "Advances in homology protein structure modeling" (<http://www.bentham-direct.org/pages/content.php?CPPS/2006/00000007/00000003/0004K.SGM>). *Current Protein and Peptide Science* **7** (3): 217–27. doi:10.2174/138920306777452312. PMID 16787261. PMC 1839925. .
- [62] Zhang Y, Skolnick J. (2005). "The protein structure prediction problem could be solved using the current PDB library" (<http://www.pnas.org/cgi/pmidlookup?view=long&pmid=15653774>). *Proceedings of the National Academy of Sciences U.S.A.* **102** (4): 1029–34. doi:10.1073/pnas.0407152101. PMID 15653774. PMC 545829. .
- [63] Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D. (2003). "Design of a novel globular protein fold with atomic-level accuracy" (<http://www.sciencemag.org/cgi/pmidlookup?view=long&pmid=14631033>). *Science* **302** (5649): 1364–68. doi:10.1126/science.1089427. PMID 14631033. .
- [64] Ritchie DW. (2008). "Recent progress and future directions in protein–protein docking" (<http://www.bentham-direct.org/pages/content.php?CPPS/2008/00000009/00000001/0001K.SGM>). *Current Protein and Peptide Science* **9** (1): 1–15. doi:10.2174/138920308783565741. PMID 18336319. .
- [65] Scheraga HA, Khalili M, Liwo A. (2007). "Protein-folding dynamics: overview of molecular simulation techniques" (http://arjournals.annualreviews.org/doi/abs/10.1146/annurev.physchem.58.032806.104614?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub=ncbi.nlm.nih.gov). *Annual Review of Physical Chemistry* **58**: 57–83. doi:10.1146/annurev.physchem.58.032806.104614. PMID 17034338. .
- [66] Zagrovic B, Snow CD, Shirts MR, Pande VS. (2002). "Simulation of folding of a small alpha-helical protein in atomistic detail using worldwide-distributed computing" (<http://linkinghub.elsevier.com/retrieve/pii/S002228360200997X>). *Journal of Molecular Biology* **323** (5): 927–37. doi:10.1016/S0022-2836(02)00997-X. PMID 12417204. . Retrieved 2009-04-14.
- [67] Herges T, Wenzel W. (2005). "In silico folding of a three helix protein and characterization of its free-energy landscape in an all-atom force field" (<http://link.aps.org/abstract/PRL/v94/p018101>). *Physical Review Letters* **94** (1): 018101. doi:10.1103/PhysRevLett.94.018101. PMID 15698135. . Retrieved 2009-04-14.
- [68] Hoffmann M, Wanko M, Strodel P, König PH, Frauenheim T, Schulten K, Thiel W, Tajkhorshid E, Elstner M. (2006). "Color tuning in rhodopsins: the mechanism for the spectral shift between bacteriorhodopsin and sensory rhodopsin II". *Journal of the American Chemical Society* **128** (33): 10808–18. doi:10.1021/ja062082i. PMID 16910676.

- [69] Brosnan J. (1 June 2003). "Interorgan amino acid transport and its regulation" (<http://jn.nutrition.org/cgi/content/full/133/6/2068S>). *Journal of Nutrition* **133** (6 Suppl 1): 2068S–72S. PMID 12771367. .
- [70] Perrett D., David (2007). "From 'protein' to the beginnings of clinical proteomics". *Proteomics — Clinical Applications* **1** (8): 720–38. doi:10.1002/prca.200700525.
- [71] New Oxford Dictionary of English
- [72] Reynolds JA, Tanford C. (2003). *Nature's Robots: A History of Proteins (Oxford Paperbacks)*. Oxford University Press, USA. p. 15. ISBN 0-19-860694-X.
- [73] Pauling L, Corey RB, Branson HR., L. (1951). "The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain" (<http://www.pnas.org/site/misc/Protein8.pdf>) (PDF). *Proceedings of the National Academy of Sciences U.S.A.* **37**: 235–40. doi:10.1073/pnas.37.5.235. .
- [74] Kauzmann W. (1956). "Structural factors in protein denaturation". *Journal of Cellular Physiology. Supplement* **47** (Suppl 1): 113–31. doi:10.1002/jcp.1030470410. PMID 13332017.
- [75] Kauzmann W. (1959). "Some factors in the interpretation of protein denaturation". *Advances in Protein Chemistry* **14**: 1–63. doi:10.1016/S0065-3233(08)60608-7. PMID 14404936.
- [76] Kalman SM, Linderstrom-Lang K, Ottesen M, Richards FM. (1955). "Degradation of ribonuclease by subtilisin". *Biochimica et Biophysica Acta* **16** (2): 297–99. doi:10.1016/0006-3002(55)90224-9. PMID 14363272.
- [77] Sanger F. (1949). "The terminal peptides of insulin" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1275055>). *Biochemical Journal* **45** (5): 563–74. PMID 15396627. PMC 1275055.
- [78] "RCSB Protein Data Bank" (<http://www.rcsb.org/pdb/home/home.do>). . Retrieved 2009-04-14.
- [79] Zhou ZH. (2008). "Towards atomic resolution structural determination by single-particle cryo-electron microscopy" ([http://linkinghub.elsevier.com/retrieve/pii/S0959-440X\(08\)00036-5](http://linkinghub.elsevier.com/retrieve/pii/S0959-440X(08)00036-5)). *Current Opinion in Structural Biology* **18** (2): 218–28. doi:10.1016/j.sbi.2008.03.004. PMID 18403197. PMC 2714865. .
- [80] Keskin O, Tuncbag N, Gursoy A. (2008). "Characterization and prediction of protein interfaces to infer protein-protein interaction networks" (<http://www.bentham-direct.org/pages/content.php?CPB/2008/00000009/00000002/0003G.SGM>). *Current Pharmaceutical Biotechnology* **9** (2): 67–76. doi:10.2174/138920108783955191. PMID 18393863. . Retrieved 2009-04-14.
- [81] <http://www.tjmitchell.com/stuart/dna.html>
- [82] <http://ctd.mdibl.org/>
- [83] <http://harvester.fzk.de>
- [84] <http://www.rcsb.org>
- [85] http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/index.html
- [86] <http://www.proteopedia.org>
- [87] <http://www.expasy.uniprot.org>
- [88] <http://www.jcvi.org/pn-utility>
- [89] <http://www.proteinatlas.org>
- [90] <http://www.ncbi.nlm.nih.gov/sites/entrez?db=protein>
- [91] <http://www.ncbi.nlm.nih.gov/sites/entrez?db=structure>
- [92] <http://www.hprd.org/>
- [93] <http://www.humanproteinpedia.org/>
- [94] <http://folding.stanford.edu/>
- [95] <http://hopes.stanford.edu/basics/proteins/p0.html>
- [96] <http://www.biochemweb.org/proteins.shtml>

Protein structure

Proteins are an important class of biological macromolecules present in all biological organisms, made up of such elements as carbon, hydrogen, nitrogen, oxygen, and sulphur. All proteins are polymers of amino acids. According to their physical size, proteins are nanoparticles (definition: 1-100 nm). The polymers, also known as polypeptides, consist of a sequence of 20 different L- α -amino acids, also referred to as residues. For chains under 40 residues the term peptide is frequently used instead of protein. To be able to perform their biological function, proteins fold into one or more specific spatial conformations, driven by a number of noncovalent interactions such as hydrogen bonding, ionic interactions, Van Der Waals forces and hydrophobic packing. To understand the functions of proteins at a molecular level, it is often necessary to determine their three dimensional structure. This is the topic of the scientific field of structural biology, that employs techniques such as X-ray crystallography, NMR spectroscopy, and Dual Polarisation Interferometry to determine the structure of proteins.

A number of residues is necessary to perform a particular biochemical function, and around 40-50 residues appears to be the lower limit for a functional domain size. Protein sizes range from this lower limit to several thousand residues in multi-functional or structural proteins. However, the current estimate for the average protein length is around 300 residues.^[1] Very large aggregates can be formed from protein subunits, for example many thousand actin molecules assemble into a microfilament.

Levels of protein structure

Biochemistry refers to four distinct aspects of a protein's structure:

Primary structure

the amino acid sequence of the peptide chains.

Secondary structure

highly regular sub-structures (*alpha helix* and *strands of beta pleated sheet*), which are locally defined, meaning that there can be many different secondary motifs present in one single protein molecule.

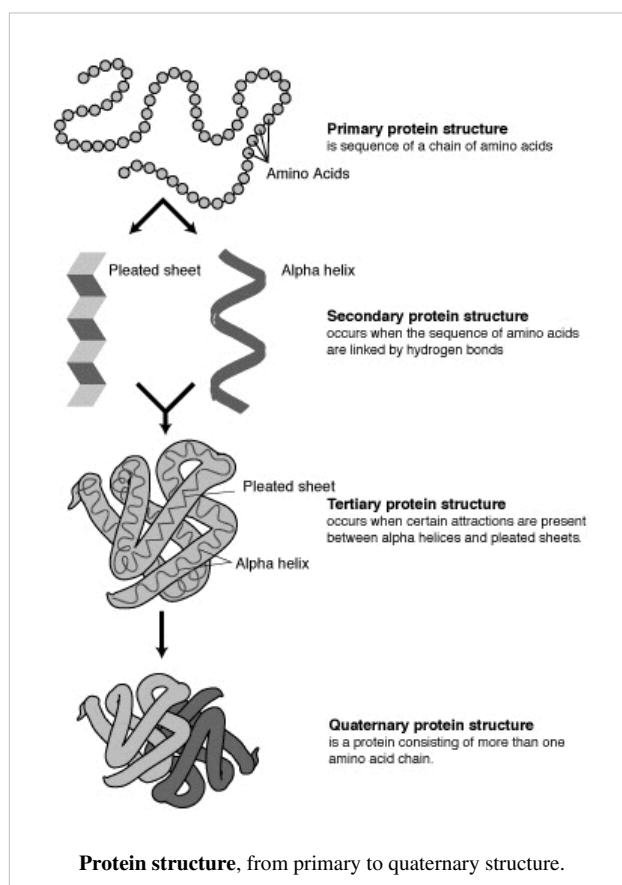
Tertiary structure

three-dimensional structure of a single protein molecule; a spatial arrangement of the secondary structures. It also describes the completely folded and compacted polypeptide chain.

Quaternary structure

complex of several protein molecules or polypeptide chains, usually called protein subunits in this context, which function as part of the larger assembly or protein complex.

In addition to these levels of structure, a protein may shift between several similar structures in performing its biological function. This process is also reversible. In the context of these functional rearrangements, these tertiary or quaternary structures are usually referred to as chemical conformation, and transitions between them are called conformational changes.



The primary structure is held together by covalent or peptide bonds, which are made during the process of protein biosynthesis or translation. These peptide bonds provide rigidity to the protein. The two ends of the amino acid chain are referred to as the C-terminal end or carboxyl terminus (C-terminus) and the N-terminal end or amino terminus (N-terminus) based on the nature of the free group on each extremity.

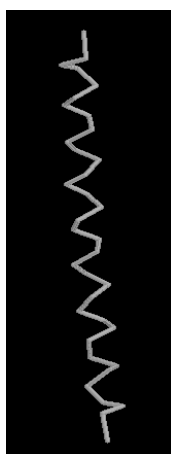
The various types of secondary structure are defined by their patterns of hydrogen bonds between the main-chain peptide groups. However, these hydrogen bonds are generally not stable by themselves, since the water-amide hydrogen bond is generally more favorable than the amide-amide hydrogen bond. Thus, secondary structure is stable only when the local concentration of water is sufficiently low, e.g., in the molten globule or fully folded states.

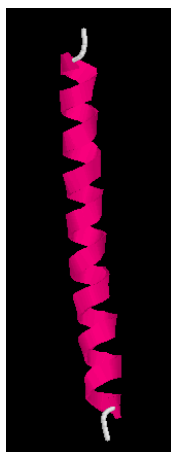
Similarly, the formation of molten globules and tertiary structure is driven mainly by structurally *non-specific* interactions, such as the rough propensities of the amino acids and hydrophobic interactions. However, the tertiary structure is *fixed* only when the parts of a protein domain are locked into place by structurally *specific* interactions, such as ionic interactions (salt bridges), hydrogen bonds, and the tight packing of side chains. The tertiary structure of extracellular proteins can also be stabilized by disulfide bonds, which reduce the entropy of the unfolded state; disulfide bonds are extremely rare in cytosolic proteins, since the cytosol is generally a reducing environment.

Primary structure

The sequence of the different amino acids is called the primary structure of the peptide or protein. Counting of residues always starts at the N-terminal end (NH_2 -group), which is the end where the amino group is involved in a peptide bond. The primary structure of a protein is determined by the gene corresponding to the protein. A specific sequence of nucleotides in DNA is transcribed into mRNA, which is read by the ribosome in a process called translation. The sequence of a protein is unique to that protein, and defines the structure and function of the protein. The sequence of a protein can be determined by methods such as Edman degradation or tandem mass spectrometry. Often however, it is read directly from the sequence of the gene using the genetic code. Post-translational modifications such as disulfide formation, phosphorylations and glycosylations are usually also considered a part of the primary structure, and cannot be read from the gene.

Secondary structure





Left: Ca atom trace. Right: Secondary structure cartoon ("ribbon")

By building models of peptides using known information about bond lengths and angles, the first elements of secondary structure, the alpha helix and the beta sheet, were suggested in 1951 by Linus Pauling and coworkers.^[2] Each of these two secondary structure elements have a regular geometry, meaning they are constrained to specific values of the dihedral angles ψ and ϕ . Thus they can be found in a specific region of the Ramachandran plot. Both the alpha helix and the beta-sheet represent a way of saturating all the hydrogen bond donors and acceptors in the peptide backbone. These secondary structure elements only depend on properties of the polypeptide *main chain*, explaining why they occur in all proteins. The part of the protein that is not in a regular secondary structure is said to be a "non-regular structure" (not to be mixed with random coil, an unfolded polypeptide chain lacking any fixed three-dimensional structure). Some more representations of the same helix are shown at right.

Supersecondary structure

The elements of secondary structure are usually folded into a compact shape using a variety of loops and turns. Secondary structures are bonded by hydrogen bond to form supersecondary structures like Greek key.^[3] See Supersecondary structure for a detailed example. It is also suggested by many scientists that the secondary structure consists only of the amino acid sequence.

Tertiary structure

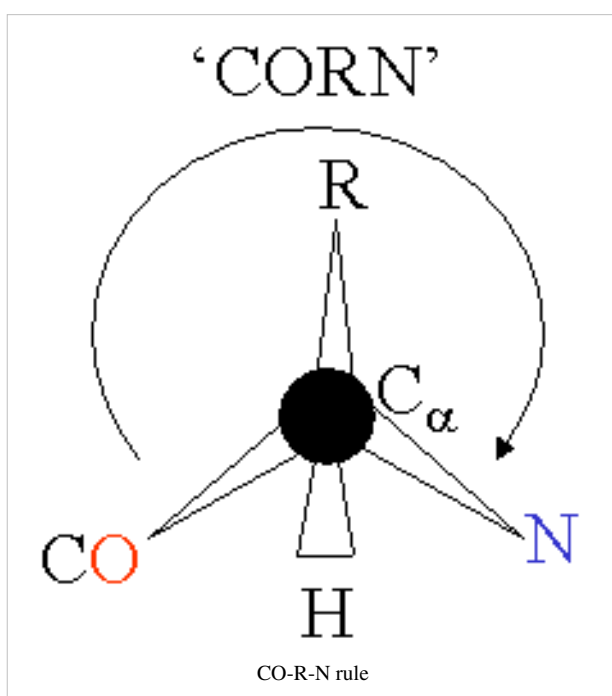
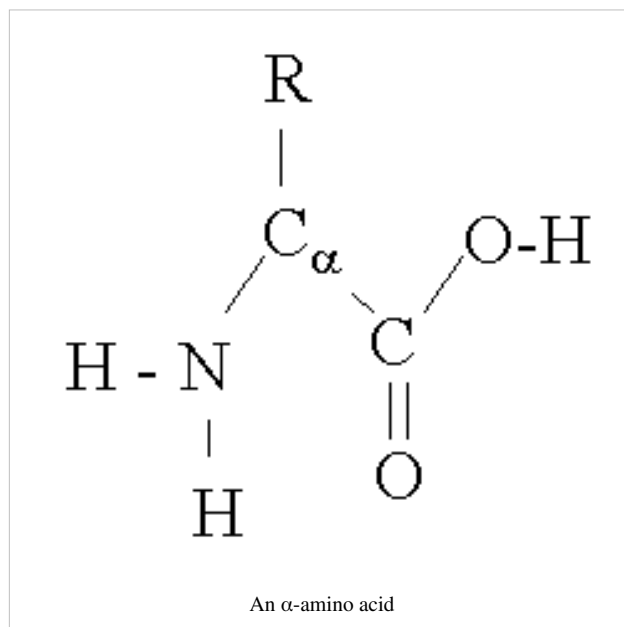
The elements of secondary structure are usually folded into a compact shape using a variety of loops and turns. The formation of tertiary structure is usually driven by the burial of hydrophobic residues, but other interactions such as hydrogen bonding, ionic interactions and disulfide bonds can also stabilize the tertiary structure. The tertiary structure encompasses all the noncovalent interactions that are not considered secondary structure, and is what defines the overall fold of the protein, and is usually indispensable for the function of the protein.

Quaternary structure

The quaternary structure is the interaction between several chains of peptide bonds. The individual chains are called subunits. The individual subunits are usually not covalently connected, but might be connected by a disulfide bond. Not all proteins have quaternary structure, since they might be functional as monomers. The quaternary structure is stabilized by the same range of interactions as the tertiary structure. Complexes of two or more polypeptides (i.e. multiple subunits) are called multimers. Specifically it would be called a dimer if it contains two subunits, a trimer if it contains three subunits, and a tetramer if it contains four subunits. The subunits are usually related to one another by symmetry axes, such as a 2-fold axis in a dimer. Multimers made up of identical subunits may be referred to with a prefix of "homo-" (e.g. a homotetramer) and those made up of different subunits may be referred to with a prefix of "hetero-" (e.g. a heterotetramer, such as the two alpha and two beta chains of hemoglobin).

Structure of the amino acids

An α -amino acid consists of a part that is present in all the amino acid types, and a side chain that is unique to each type of residue. The C_{α} atom is bound to 4 different atoms: a hydrogen atom (the H is omitted in the diagram), an amino group nitrogen, a carboxyl group carbon, and a side chain carbon specific for this type of amino acid. An exception from this rule is proline, where the hydrogen atom is replaced by a bond to the side chain. Because the carbon atom is bound to four different groups it is chiral, however only one of the isomers occur in biological proteins. Glycine however, is not chiral since its side chain is a hydrogen atom. A simple mnemonic for correct L-form is "CORN": when the C_{α} atom is viewed with the H in front, the residues read "CO-R-N" in a clockwise direction.



The side chain determines the chemical properties of the α -amino acid and may be any one of the 20 different side chains:

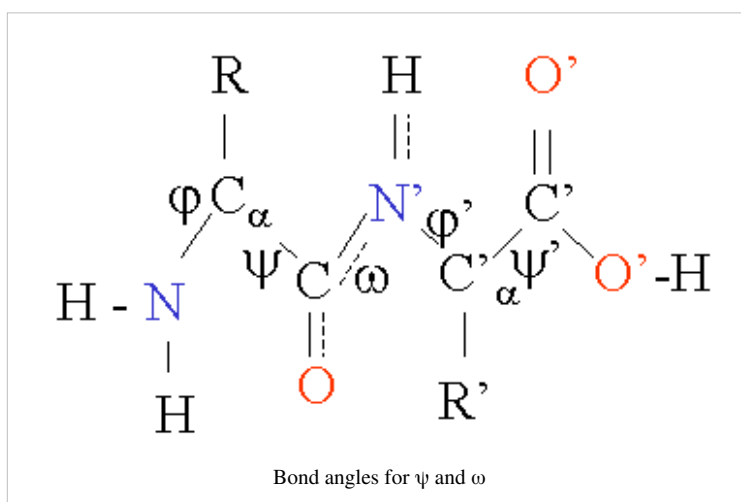
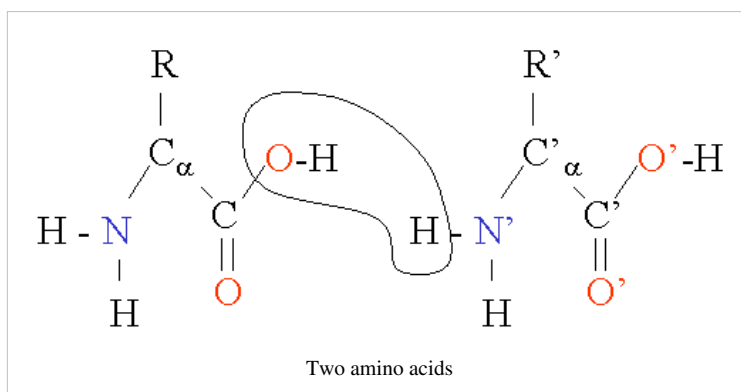
The 20 naturally occurring amino acids can be divided into several groups based on their chemical properties. Important factors are charge, hydrophobicity/hydrophilicity, size and functional groups. The nature of the interaction of the different side chains with the aqueous environment plays a major role in molding protein structure. Hydrophobic side chains tends to be buried in the middle of the protein, whereas hydrophilic side chains are exposed to the solvent.

Examples of hydrophobic residues are: Leucine, isoleucine, phenylalanine, and valine, and to a lesser extent tyrosine, alanine and tryptophan. The charge of the side chains plays an important role in protein structures, since ion bonding can stabilize proteins structures, and an unpaired charge in the middle of a

protein can disrupt structures. Charged residues are strongly hydrophilic, and are usually found on the out side of proteins. Positively charged side chains are found in lysine and arginine, and in some cases in histidine. Negative charges are found in glutamate and aspartate. The rest of the amino acids have smaller generally hydrophilic side chains with various functional groups. Serine and threonine have hydroxyl groups, and asparagine and glutamine have amide groups. Some amino acids have special properties such as cysteine, that can form covalent disulfide bonds to other cysteines, proline that is cyclical, and glycine that is small, and more flexible than the other amino acids.

The peptide bond (*amide bond*)

Two amino acids can be combined in a condensation reaction. By repeating this reaction, long chains of residues (amino acids in a peptide bond) can be generated. This reaction is catalysed by the ribosome in a process known as translation. The peptide bond is in fact planar due to the delocalization of the electrons from the double bond. The rigid peptide dihedral angle, ω (the bond between C_1 and N) is always close to 180 degrees. The dihedral angles phi ϕ (the bond between N and C_α) and psi ψ (the bond between C_α and C_1) can have a certain range of possible values. These angles are the degrees of freedom of a protein, they control the protein's three dimensional structure. They are restrained by geometry to allowed ranges typical for particular secondary structure elements, and represented in a Ramachandran plot. A few important bond lengths are given in the table below.



Peptide bond	Average length	Single bond	Average length	Hydrogen bond	Average (± 30)
Ca - C	153 pm	C - C	154 pm	O-H --- O-H	280 pm
C - N	133 pm	C - N	148 pm	N-H --- O=C	290 pm
N - Ca	146 pm	C - O	143 pm	O-H --- O=C	280 pm

Side-chain conformation and Rotamers

The atoms along the side chain are named with Greek letters in Greek alphabetical order: α , β , γ , δ , ϵ , and so on. C_α refers to the carbon atom of the backbone closest to the carbonyl group of that amino acid, C_β the second closest and so on. The C_α is part of the backbone, while C_β and atoms further out comprise the side chain. The dihedral angles around the bonds between these atoms are named χ_1 , χ_2 , χ_3 , etc. The dihedral angle of the first movable atom of the side chain, γ , defined as N-C α -C β -X γ , is named χ_1 . Most side chains can be in different conformations called gauche(-), trans, and gauche(+). Side chains generally tend to try to come into a staggered conformation around χ_2 , driven by the minimization of the overlap between the electron orbitals of substituent atoms.

The diversity of side-chain conformations is often expressed in rotamer libraries. A rotamer library is a collection of rotamers for each residue type in proteins with side-chain degrees of freedom. Rotamer libraries usually contain information about both conformation and frequency of a certain conformation. Often libraries will also contain information about the variance about dihedral angle means or modes, which can be used in sampling.^[4]

Side-chain dihedral angles are not evenly distributed, but for most side chain types, the χ angles occur in tight clusters around certain values. Rotamer libraries therefore are usually derived from statistical analysis of side-chain conformations in known structures of proteins by clustering observed conformations or by dividing dihedral angle space into bins, and determining an average conformation in each bin. This division is usually on physical-chemical grounds, as in the divisions for rotation about sp³-sp³ bonds into three 120° bins centered on each staggered conformation (60°, 180°, -60°).

Rotamer libraries can be backbone-independent, secondary-structure-dependent, or backbone-dependent. The distinctions are made depending on whether the dihedral angles for the rotamers and/or their frequencies depend on the local backbone conformation or not. Backbone-independent rotamer libraries make no reference to backbone conformation, and are calculated from all available side chains of a certain type.^[5] Secondary-structure-dependent libraries present different dihedral angles and/or rotamer frequencies for α -helix, β -sheet, or coil secondary structures. Backbone-dependent rotamer libraries present conformations and/or frequencies dependent on the local backbone conformation as defined by the backbone dihedral angles ϕ and ψ , regardless of secondary structure.^[6] Finally, a variant on backbone-dependent rotamer libraries exists in the form of position-specific rotamers, those defined by a fragment usually of 5 amino acids in length, where the central residue's side chain conformation is examined.

Domains, motifs, and folds in protein structure

Many proteins are organized into several units. A structural domain is an element of the protein's overall structure that is self-stabilizing and often folds independently of the rest of the protein chain. Many domains are not unique to the protein products of one gene or one gene family but instead appear in a variety of proteins. Domains often are named and singled out because they figure prominently in the biological function of the protein they belong to; for example, the "calcium-binding domain of calmodulin". Because they are self-stabilizing, domains can be "swapped" by genetic engineering between one protein and another to make chimeras. A motif in this sense refers to a small specific combination of secondary structural elements (such as helix-turn-helix). These elements are often called supersecondary structures. Fold refers to a global type of arrangement, like helix bundle or beta-barrel. Structure motifs usually consist of just a few elements, e.g. the 'helix-turn-helix' has just three. Note that while the *spatial sequence* of elements is the same in all instances of a motif, they may be encoded in any order within the underlying gene. Protein structural motifs often include loops of variable length and unspecified structure, which in effect create the "slack" necessary to bring together in space two elements that are not encoded by immediately adjacent DNA sequences in a gene. Note also that even when two genes encode secondary structural elements of a motif in the same order, nevertheless they may specify somewhat different sequences of amino acids. This is true not only because of the complicated relationship between tertiary and primary structure, but because the size of the elements varies from one protein and the next. Despite the fact that there are about 100,000 different proteins expressed in eukaryotic systems, there are much fewer different domains, structural motifs and folds. This is partly a consequence of evolution, since genes or parts of genes can be doubled or moved around within the genome. This means that, for example, a protein domain might be moved from one protein to another thus giving the protein a new function. Because of these mechanisms pathways and mechanisms tends to be reused in several different proteins.

Protein folding

An unfolded polypeptide folds into its characteristic three-dimensional structure from random coil.

Protein structure determination

Around 90% of the protein structures available in the Protein Data Bank have been determined by X-ray crystallography. This method allows one to measure the 3D density distribution of electrons in the protein (in the crystallized state) and thereby infer the 3D coordinates of all the atoms to be determined to a certain resolution. Roughly 9% of the known protein structures have been obtained by Nuclear Magnetic Resonance techniques, which can also be used to determine secondary structure. Note that aspects of the secondary structure as whole can be determined via other biochemical techniques such as circular dichroism or dual polarisation interferometry. Secondary structure can also be predicted with a high degree of accuracy (see next section). Cryo-electron microscopy has recently become a means of determining protein structures to high resolution (less than 5 angstroms or 0.5 nanometer) and is anticipated to increase in power as a tool for high resolution work in the next decade. This technique is still a valuable resource for researchers working with very large protein complexes such as virus coat proteins and amyloid fibers.

A rough guide to the resolution of protein structures

Resolution (Å)	Meaning
>4.0	Individual coordinates meaningless
3.0 - 4.0	Fold possibly correct, but errors are very likely. Many sidechains placed with wrong rotamer.
2.5 - 3.0	Fold likely correct except that some surface loops might be mismodelled. Several long, thin sidechains (lys, glu, gln, etc) and small sidechains (ser, val, thr, etc) likely to have wrong rotamers.
2.0 - 2.5	As 2.5 - 3.0, but number of sidechains in wrong rotamer is considerably less. Many small errors can normally be detected. Fold normally correct and number of errors in surface loops is small. Water molecules and small ligands become visible.
1.5 - 2.0	Few residues have wrong rotamer. Many small errors can normally be detected. Folds are extremely rarely incorrect, even in surface loops.
0.5 - 1.5	In general, structures have almost no errors at this resolution. Rotamer libraries and geometry studies are made from these structures.

Structure classification

Protein structures can be classified based on their similarity or a common evolutionary origin. SCOP and CATH databases provide two different structural classifications of proteins.

Computational prediction of protein structure

The generation of a protein sequence is much simpler than the generation of a protein structure. However, the structure of a protein gives much more insight in the function of the protein than its sequence. Therefore, a number of methods for the computational prediction of protein structure from its sequence have been proposed. *Ab initio* prediction methods use just the sequence of the protein. Threading uses existing protein structures. Homology Modeling to build a reliable 3D model for a protein of unknown structure from one or more related proteins of known structure. The recent progress and challenges in protein structure prediction was reviewed by Zhang^[7].

Protein structure related software

There are software to aid researchers working on, often overlapping, different aspects of protein structure. The most basic functionality is providing structure visualization. Analysis of protein structure can be facilitated by software that aligns structures. In the absence of existing structures for a given protein sequence, there are methods to predict or to model the structure of such sequences based on known protein structures. And given models of known or predicted structures, one can use software to verify them for errors, predict protein conformational changes, or predict substrate binding sites.

See also

- Protein dynamics

Further reading

- Chiang YS, Gelfand TI, Kister AE, Gelfand IM (2007). "New classification of supersecondary structures of sandwich-like proteins uncovers strict patterns of strand assemblage.". *Proteins*. **68** (4): 915–921. doi:10.1002/prot.21473. PMID 17557333.
- Habeck M, Nilges M, Rieping W (2005). "Bayesian inference applied to macromolecular structure determination"^[8]. *Physical review. E, Statistical, nonlinear, and soft matter physics* **72** (3 Pt 1): 031912. PMID 16241487. (Bayesian computational methods for the structure determination from NMR data)

External links

- SSS Database^[9] — super-secondary structure protein database
- SPROUTS^[10] (Structural Prediction for pRotein fOlding UTility System)
- ProSA-web^[11] — a web service for the recognition of errors in experimentally or theoretically determined protein structures
- NQ-Flipper^[12] — checks for unfavorable rotamers of Asn and Gln residues in protein structures
- WHAT IF servers^[13] — checks nearly 200 aspects of protein structure, like packing, geometry, unfavourable rotamers in general of for Asn, Gln, and His especially, strange water molecules, backbone conformations, atom nomenclature, symmetry parameters, etc.
- Bioinformatics course^[14] — an interactive, fully free, course explaining many of the aspects discussed in this wiki entry.

References

- [1] Brocchieri L, Karlin S (2005-06-10). "Protein length in eukaryotic and prokaryotic proteomes" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1150220>). *Nucleic Acids Research* **33** (10): 3390–3400. doi:10.1093/nar/gki615. PMID 15951512. PMC 1150220.
- [2] Pauling L, Corey RB, Branson HR (1951). "The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1063337>). *Proc Natl Acad Sci USA* **37** (4): 205–211. doi:10.1073/pnas.37.4.205. PMID 14816373. PMC 1063337.
- [3] Chiang YS, Gelfand TI, Kister AE, Gelfand IM (2007). "New classification of supersecondary structures of sandwich-like proteins uncovers strict patterns of strand assemblage.". *Proteins*. **68** (4): 915–921. doi:10.1002/prot.21473. PMID 17557333.
- [4] Dunbrack, RL (2002). "Rotamer Libraries in the 21st Century". *Curr Opin. Struct. Biol.* **12** (4): 431–440. doi:10.1016/S0959-440X(02)00344-5. PMID 12163064.
- [5] Richardson Rotamer Libraries (<http://pibs.duke.edu/databases/rotamer.php>)
- [6] Dunbrack Rotamer Libraries (<http://dunbrack.fccc.edu/bbdep>)
- [7] Zhang Y (2008). "Progress and challenges in protein structure prediction" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2680823>). *Curr Opin Struct Biol* **18** (3): 342–348. doi:10.1016/j.sbi.2008.02.004. Entrez Pubmed 18436442 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=18436442). PMID 18436442. PMC 2680823.

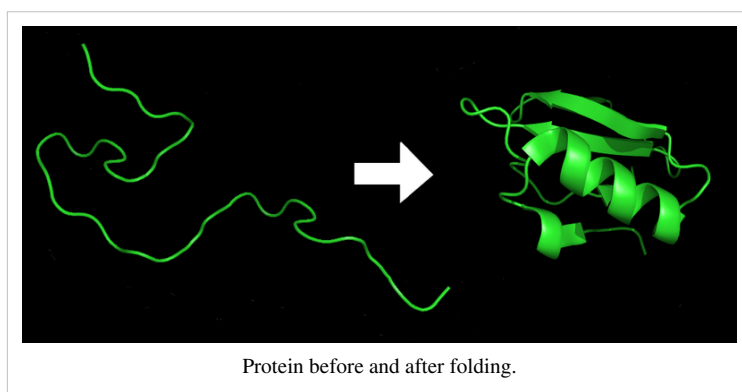
- [8] <http://www.spineurope.org/publications/Habeck%20et%20al%20031912%202005.pdf>
- [9] <http://binfs.umdj.edu/sssd/>
- [10] <http://bioinformatics.eas.asu.edu/springs/Sprouts/projectsSprouts.html>
- [11] <https://prosa.services.came.sbg.ac.at/prosa.php>
- [12] <https://flipper.services.came.sbg.ac.at/>
- [13] <http://swift.cmbi.ru.nl/>
- [14] <http://swift.cmbi.ru.nl/teach/B1/>

Protein folding

Protein folding is the physical process by which a polypeptide folds into its characteristic and functional three-dimensional structure from random coil.^[1] Each protein exists as an unfolded polypeptide or random coil when translated from a sequence of mRNA to a linear chain of amino acids. This polypeptide lacks any developed three-dimensional structure (the left hand side of the neighboring figure). Amino acids interact with each other to

produce a well-defined three dimensional structure, the folded protein (the right hand side of the figure), known as the native state. The resulting three-dimensional structure is determined by the amino acid sequence.^[2]

For many proteins the correct three dimensional structure is essential to function.^[3] Failure to fold into the intended shape usually produces inactive proteins with different properties including toxic prions. Several neurodegenerative and other diseases are believed to result from the accumulation of *misfolded* (incorrectly folded) proteins.^[4]



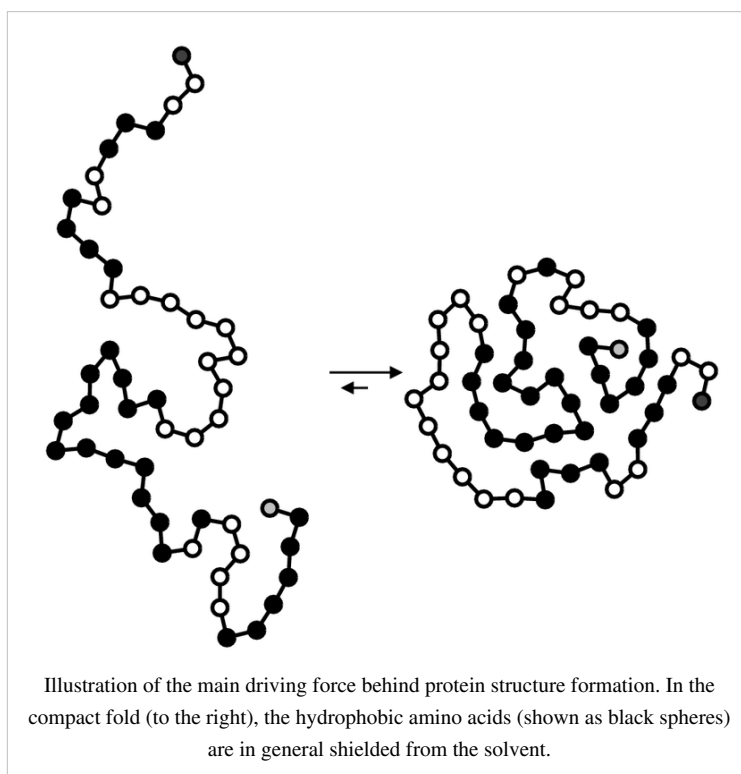
Known facts

Relationship between folding and amino acid sequence

The amino-acid sequence (or primary structure) of a protein defines its native conformation. A protein molecule folds spontaneously during or after synthesis. While these macromolecules may be regarded as "folding themselves", the process also depends on the solvent (water or lipid bilayer),^[5] the concentration of salts, the temperature, and the presence of molecular chaperones.

Folded proteins usually have a hydrophobic core in which side chain packing stabilizes the folded state, and charged or polar side chains occupy the solvent-exposed surface where they interact with surrounding water. Minimizing the number of hydrophobic side-chains exposed to water is an important driving force behind the folding process.^[6]

Formation of intramolecular hydrogen bonds provides another important contribution to protein stability.^[7] The strength of hydrogen bonds depends on their environment, thus H-bonds enveloped in a hydrophobic core contribute more than H-bonds exposed to the aqueous environment to the stability of the native state.^[8]



In the seminal research work published nearly four decades ago, C.B. Anfinsen hypothesized that "information dictating the native fold of protein domains is encoded in their amino acid sequence"^[9]. However, with the explosive amount of protein sequence, structure, and fold data generated since the time of Anfinsen during the omics era, the emerging picture of the protein universe has challenged Anfinsen's dogma, for it has become evident that numerous protein folds have incredible sequence diversity with no consistent "fold code"^[10]. In support of this observation, recent studies have shown that proteins with as low as 1-2% sequence identity may still adopt the same native fold, thus defying any tangible encoding of fold-dictating information into protein sequence^[11]. The pursuit of the elusive "fold code" has resulted in little more than patterns of amino acid sequence conservations specific to certain proteins, but no finding has been compelling enough to generalize universally or to utilize for biological applications^[12].

In a recent study, scientists from Harvard-MIT have shown that, despite the enormous diversity within protein folds at the level of 1-dimensional amino acid sequence, nature has encoded fold-conserved information at higher dimensions of protein space such as the 2-D (protein contact maps) or 3-D (structure), that are known to be more intricately related to protein folding phenomena^[13]. The study published in PLoS ONE illuminated latent fold-conserved information from higher dimensional protein space using network theory approaches. By examining the entire protein universe on a fold-by-fold basis, the study revealed that atomic interaction networks in the solvent-unexposed core of protein domains are fold-conserved and unique to each protein's native fold, thus appearing to be the encoded "signature" of protein domains. This study hence uncovered that the protein fold code is a "network phenomena" in addition to a sequence and structural phenomena as commonly presumed. The discovery of such a protein folding code also confirms Anfinsen's Dogma by proving that a significant portion of the fold-dictating

information is encoded by the atomic interaction network in the solvent-unexposed core of protein domains.

The process of folding *in vivo* often begins co-translationally, so that the N-terminus of the protein begins to fold while the C-terminal portion of the protein is still being synthesized by the ribosome. Specialized proteins called chaperones assist in the folding of other proteins.^[14] A well studied example is the bacterial GroEL system, which assists in the folding of globular proteins. In eukaryotic organisms chaperones are known as heat shock proteins. Although most globular proteins are able to assume their native state unassisted, chaperone-assisted folding is often necessary in the crowded intracellular environment to prevent aggregation; chaperones are also used to prevent misfolding and aggregation which may occur as a consequence of exposure to heat or other changes in the cellular environment.

For the most part, scientists have been able to study many identical molecules folding together *en masse*. At the coarsest level, it appears that in transitioning to the native state, a given amino acid sequence takes on roughly the same route and proceeds through roughly the same intermediates and transition states. Often folding involves first the establishment of regular secondary and supersecondary structures, particularly alpha helices and beta sheets, and afterwards tertiary structure. Formation of quaternary structure usually involves the "assembly" or "coassembly" of subunits that have already folded. The regular alpha helix and beta sheet structures fold rapidly because they are stabilized by intramolecular hydrogen bonds, as was first characterized by Linus Pauling. Protein folding may involve covalent bonding in the form of disulfide bridges formed between two cysteine residues or the formation of metal clusters. Shortly before settling into their more energetically favourable native conformation, molecules may pass through an intermediate "molten globule" state.

The essential fact of folding, however, remains that the amino acid sequence of each protein contains the information that specifies both the native structure and the pathway to attain that state. This is not to say that nearly identical amino acid sequences always fold similarly.^[15] Conformations differ based on environmental factors as well; similar proteins fold differently based on where they are found. Folding is a spontaneous process independent of energy inputs from nucleoside triphosphates. The passage of the folded state is mainly guided by hydrophobic interactions, formation of intramolecular hydrogen bonds, and van der Waals forces, and it is opposed by conformational entropy.

Disruption of the native state

Under some conditions proteins will not fold into their biochemically functional forms. Temperatures above or below the range that cells tend to live in will cause thermally unstable proteins to unfold or "denature" (this is why boiling makes an egg white turn opaque). High concentrations of solutes, extremes of pH, mechanical forces, and the presence of chemical denaturants can do the same. Protein thermal stability is far from constant, however. For example, hyperthermophilic bacteria have been found that grow at temperatures as high as 122°C,^[16] which of course requires that their full complement of vital proteins and protein assemblies be stable at that temperature or above.

A fully denatured protein lacks both tertiary and secondary structure, and exists as a so-called random coil. Under certain conditions some proteins can refold; however, in many cases denaturation is irreversible.^[17] Cells sometimes protect their proteins against the denaturing influence of heat with enzymes known as chaperones or heat shock proteins, which assist other proteins both in folding and in remaining folded. Some proteins never fold in cells at all except with the assistance of chaperone molecules, which either isolate individual proteins so that their folding is not interrupted by interactions with other proteins or help to unfold misfolded proteins, giving them a second chance to refold properly. This function is crucial to prevent the risk of precipitation into insoluble amorphous aggregates.

Incorrect protein folding and neurodegenerative disease

Aggregated proteins are associated with prion-related illnesses such as Creutzfeldt-Jakob disease, bovine spongiform encephalopathy (mad cow disease), amyloid-related illnesses such as Alzheimer's Disease and familial amyloid cardiomyopathy or polyneuropathy, as well as intracytoplasmic aggregation diseases such as Huntington's and Parkinson's disease.^{[4] [18]} These age onset degenerative diseases are associated with the multimerization of misfolded proteins into insoluble, extracellular aggregates and/or intracellular inclusions including cross-beta sheet amyloid fibrils; it is not clear whether the aggregates are the cause or merely a reflection of the loss of protein homeostasis, the balance between synthesis, folding, aggregation and protein turnover. Misfolding and excessive degradation instead of folding and function leads to a number of proteopathy diseases such as antitrypsin-associated Emphysema, cystic fibrosis and the lysosomal storage diseases, where loss of function is the origin of the disorder. While protein replacement therapy has historically been used to correct the latter disorders, an emerging approach is to use pharmaceutical chaperones to fold mutated proteins to render them functional.

The Levinthal paradox and Kinetics

The Levinthal paradox^[19] observes that if a protein were to fold by sequentially sampling all possible conformations, it would take an astronomical amount of time to do so, even if the conformations were sampled at a rapid rate (on the nanosecond or picosecond scale). Based upon the observation that proteins fold much faster than this, Levinthal then proposed that a random conformational search does not occur, and the protein must, therefore, fold through a series of meta-stable intermediate states.

The duration of the folding process varies dramatically depending on the protein of interest. When studied outside the cell, the slowest folding proteins require many minutes or hours to fold primarily due to proline isomerization, and must pass through a number of intermediate states, like checkpoints, before the process is complete.^[20] On the other hand, very small single-domain proteins with lengths of up to a hundred amino acids typically fold in a single step.^[21] Time scales of milliseconds are the norm and the very fastest known protein folding reactions are complete within a few microseconds.^[22]

Energy landscape theory of protein folding

The protein folding phenomenon was largely an experimental endeavor until the formulation of energy landscape theory by Joseph Bryngelson and Peter Wolynes in the late 1980s and early 1990s. This approach introduced the principle of minimal frustration, which asserts that evolution has selected the amino acid sequences of natural proteins so that interactions between side chains largely favor the molecule's acquisition of the folded state. Interactions that do not favor folding are selected against, although some residual *frustration* is expected to exist. A consequence of these evolutionarily selected sequences is that proteins are generally thought to have globally "funneled energy landscapes" (coined by José Onuchic[reference needed]) that are largely directed towards the native state. This "folding funnel" landscape allows the protein to fold to the native state through any of a large number of pathways and intermediates, rather than being restricted to a single mechanism. The theory is supported by both computational simulations of model proteins and numerous experimental studies, and it has been used to improve methods for protein structure prediction and design [reference needed]. The description of protein folding by the leveling free-energy landscape is also consistent with the 2nd law of thermodynamics.^[23]

Techniques for studying protein folding

Circular dichroism

Circular dichroism is one of the most general and basic tools to study protein folding. Circular dichroism spectroscopy measures the absorption of circularly polarized light. In proteins, structures such as alpha helices and beta sheets are chiral, and thus absorb such light. The absorption of this light acts as a marker of the degree of foldedness of the protein ensemble. This technique can be used to measure equilibrium unfolding of the protein by measuring the change in this absorption as a function of denaturant concentration or temperature. A denaturant melt measures the free energy of unfolding as well as the protein's m value, or denaturant dependence. A temperature melt measures the melting temperature (T_m) of the protein. This type of spectroscopy can also be combined with fast-mixing devices, such as stopped flow, to measure protein folding kinetics and to generate chevron plots.

Vibrational circular dichroism of proteins

The more recent developments of vibrational circular dichroism (VCD) techniques for proteins, currently involving Fourier transform (FFT) instruments, provide powerful means for determining protein conformations in solution even for very large protein molecules. Such VCD studies of proteins are often combined with X-ray diffraction of protein crystals, FT-IR data for protein solutions in heavy water (D_2O), or *ab initio* quantum computations to provide unambiguous structural assignments that are unobtainable from CD.

Modern studies of folding with high time resolution

The study of protein folding has been greatly advanced in recent years by the development of fast, time-resolved techniques. These are experimental methods for rapidly triggering the folding of a sample of unfolded protein, and then observing the resulting dynamics. Fast techniques in widespread use include neutron scattering^[24], ultrafast mixing of solutions, photochemical methods, and laser temperature jump spectroscopy. Among the many scientists who have contributed to the development of these techniques are Jeremy Cook, Heinrich Roder, Harry Gray, Martin Gruebele, Brian Dyer, William Eaton, Sheena Radford, Chris Dobson, Sir Alan R. Fersht and Bengt Nölting.

Computational prediction of protein tertiary structure

De novo or *ab initio* techniques for computational protein structure prediction is related to, but strictly distinct from, studies involving protein folding. Molecular Dynamics (MD) is an important tool for studying protein folding and dynamics in silico. Because of computational cost, *ab initio* MD folding simulations with explicit water are limited to peptides and very small proteins^{[25] [26]}. MD simulations of larger proteins remain restricted to dynamics of the experimental structure or its high-temperature unfolding. In order to simulate long time folding processes (beyond about 1 microsecond), like folding of small-size proteins (about 50 residues) or larger, some approximations or simplifications in protein models need to be introduced. An approach using reduced protein representation (pseudo-atoms representing groups of atoms are defined) and statistical potential is not only useful in protein structure prediction, but is also capable of reproducing the folding pathways.^[27]

There are distributed computing projects which use idle CPU or GPU time of personal computers to solve problems such as protein folding or prediction of protein structure. People can run these programs on their computer or PlayStation 3 to support them. See links below (for example Folding@Home) to get information about how to participate in these projects.

Experimental techniques of protein structure determination

Folded structures of proteins are routinely determined by X-ray crystallography and NMR.

See also

- Anfinsen's dogma
- Harvard-MIT Scientists Discover the Protein Folding Code and demonstrate this is a Network phenomenon ^[28]
- Chevron plot
- Denaturation (biochemistry)
- Denaturation midpoint
- Downhill folding
- Folding (chemistry)
- Folding@Home
- Foldit computer game
- Levinthal paradox
- Protein design
- Protein dynamics
- Protein Misfolding Cyclic Amplification
- Protein structure prediction
- Protein structure prediction software
- Rosetta@home
- Software for molecular mechanics modeling

External links

- FoldIt - Folding Protein Game ^[29]
- Folding@Home ^[30]
- Rosetta@Home ^[31]

References

- [1] Alberts, Bruce; Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walters (2002). "The Shape and Structure of Proteins" ([http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=books&doptcmdl=GenBookHL&term=mboc4\[book\]+AND+372270\[uid\]&rid=mboc4.section.388](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=books&doptcmdl=GenBookHL&term=mboc4[book]+AND+372270[uid]&rid=mboc4.section.388)). *Molecular Biology of the Cell; Fourth Edition*. New York and London: Garland Science. ISBN 0-8153-3218-1. .
- [2] Anfinsen, C. (1972). "The formation and stabilization of protein structure" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1173893>). *Biochem. J.* **128** (4): 737–49. PMID 4565129. PMC 1173893.
- [3] Jeremy M. Berg, John L. Tymoczko, Lubert Stryer; Web content by Neil D. Clarke (2002). "3. Protein Structure and Function" ([http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=books&doptcmdl=GenBookHL&term=stryer\[book\]+AND+215168\[uid\]&rid=stryer.chapter.280](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Search&db=books&doptcmdl=GenBookHL&term=stryer[book]+AND+215168[uid]&rid=stryer.chapter.280)). *Biochemistry*. San Francisco: W. H. Freeman. ISBN 0-7167-4684-0. .
- [4] Dennis J. Selkoe (2003). "Folding proteins in fatal ways" (<http://www.nature.com/nature/journal/v426/n6968/full/nature02264.html>). *Nature* **426**: pp. 900–904. doi:10.1038/nature02264. PMID 14685251. .
- [5] van den Berg, B., Wain, R., Dobson, C. M., Ellis R. J. (August 2000). "Macromolecular crowding perturbs protein refolding kinetics: implications for folding inside the cell" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=306593>). *EMBO J.* **19** (15): 3870–5. doi:10.1093/emboj/19.15.3870. PMID 10921869. PMC 306593.
- [6] Pace, C., Shirley, B., McNutt, M., Gajiwala, K. (1 January 1996). "Forces contributing to the conformational stability of proteins" (<http://www.fasebj.org/cgi/reprint/10/1/75>). *FASEB J.* **10** (1): 75–83. PMID 8566551. .
- [7] Rose, G., Fleming, P., Banavar, J., Maritan, A. (2006). "A backbone-based theory of protein folding" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1636505>). *Proc. Natl. Acad. Sci. U.S.A.* **103** (45): 16623–33. doi:10.1073/pnas.0606843103. PMID 17075053. PMC 1636505.
- [8] Deechongkit, S., Nguyen, H., Dawson, P. E., Gruebele, M., Kelly, J. W. (2004). "Context Dependent Contributions of Backbone H-Bonding to β -Sheet Folding Energetics" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1636505>). *Nature* **403** (45): 101–5. doi:10.1073/pnas.0606843103. PMID 17075053. PMC 1636505.

- [9] Anfinsen CB. (20 July 1973). "Principles that Govern the Folding of Protein Chains" (http://www.sciencemag.org/cgi/pdf_extract/181/4096/223). *Science*. **181** (96): 223–230. doi:10.1126/science.181.4096.223. PMID 4124164. .
- [10] Govindarajan S, Recabarren R, Goldstein RA. (17 Sep 1999). "Estimating the total number of protein folds." (<http://www3.interscience.wiley.com/journal/65000323/abstract>). *Proteins*. **35** (4): 408–414. doi:10.1002/(SICI)1097-0134(19990601)35:4<408::AID-PROT4>3.0.CO;2-A. PMID 10382668. .
- [11] Mirny, L. A., Abkevich, V. I. & Shakhnovich, E. I. (28 Apr 1998). "How evolution makes proteins fold quickly." (<http://www.pnas.org/content/95/9/4976.abstract>). *Proc Natl Acad Sci U S A*. **95** (9): 4976–4981. doi:10.1073/pnas.95.9.4976. PMID 9560213. PMC 20198. .
- [12] S Rackovsky. (15 Jan 1993). "On the nature of the protein folding code." (<http://www.pnas.org/content/90/2/644.abstract>). *Proc Natl Acad Sci U S A*. **90** (2): 644–648. doi:10.1073/pnas.90.2.644. PMID 8421700. PMC 45720. .
- [13] Venkataramanan Soundararajan, Rahul Raman, S. Raguram, V. Sasisekharan, Ram Sasisekharan (2010). "Atomic Interaction Networks in the Core of Protein Domains and Their Native Folds" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2826414>). *PLoS ONE* **5** (2): e9391. doi:10.1371/journal.pone.0009391. PMID 20186337. PMC 2826414.
- [14] Lee, S., Tsai, F. (2005). "Molecular chaperones in protein quality control" (<http://www.jbmb.or.kr/fulltext/jbmb/view.php?vol=38&page=259>). *J. Biochem. Mol. Biol.* **38** (3): 259–65. PMID 15943899. .
- [15] Alexander, P. A., He Y., Chen, Y., Orban, J., Bryan, P. N. (2007). "The design and characterization of two proteins with 88% sequence identity but different structure and function" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1906725>). *Proc Natl Acad Sci U S A*. **104** (29): 11963–8. doi:10.1073/pnas.0700922104. PMID 17609385. PMC 1906725.
- [16] Takai, K., Nakamura, K., Toki, T., Tsunogai, U., Miyazaki, M., Miyazaki, J., Hirayama, H., Nakagawa, S., Nunoura, T., Horikoshi, K. (2008). "Cell proliferation at 122°C and isotopically heavy CH₄ production by a hyperthermophilic methanogen under high-pressure cultivation" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2490668>). *Proc Natl Acad Sci USA* **105** (31): 10949–54. doi:10.1073/pnas.0712334105. PMID 18664583. PMC 2490668.
- [17] Shortle, D. (1 January 1996). "The denatured state (the other half of the folding equation) and its role in protein stability" (<http://www.fasebj.org/cgi/reprint/10/1/27>). *FASEB J.* **10** (1): 27–34. PMID 8566543. .
- [18] Chiti, F.; Dobson, C. (2006). "Protein misfolding, functional amyloid, and human disease." *Annual review of biochemistry* **75**: 333–366. doi:10.1146/annurev.biochem.75.101304.123901. PMID 16756495.
- [19] C. Levinthal (1968). "Are there pathways for protein folding?" (<http://www.biochem.wisc.edu/courses/biochem704/Reading/Levinthal1968.pdf>). *J. Chim. Phys.* **65**: 44–5. .
- [20] Kim, P. S., Baldwin, R. L. (1990). "Intermediates in the folding reactions of small proteins". *Annu. Rev. Biochem.* **59**: 631–60. doi:10.1146/annurev.bi.59.070190.003215. PMID 2197986.
- [21] Jackson S. E. (August 1998). "How do small single-domain proteins fold?" (<http://biomednet.com/elecref/13590278003R0081>). *Fold Des* **3** (4): R81–91. doi:10.1016/S1359-0278(98)00033-9. PMID 9710577. .
- [22] Kubelka, J., Hofrichter, J., Eaton, W. A. (February 2004). "The protein folding 'speed limit'". *Curr. Opin. Struct. Biol.* **14** (1): 76–88. doi:10.1016/j.sbi.2004.01.013. PMID 15102453.
- [23] Sharma, V., Kaila, V. R. L., and Annala, A. (2009). "Protein folding as an evolutionary process". *Physica A* **388** (6): 851–862. doi:10.1016/j.physa.2008.12.004.
- [24] Bu, Z.; Cook, J.; Callaway, D. J. E. (2001). "Dynamic regimes and correlated structural dynamics in native and denatured alpha-lactalbuminC". *J Mol Biol* **312** (4): 865–873. doi:10.1006/jmbi.2001.5006. PMID 11575938.
- [25] "Fragment-based Protein Folding Simulations" (<http://www.cs.ucl.ac.uk/staff/d.jones/t42morph.html>). .
- [26] "Protein folding" (<http://www.biomolecular-modeling.com/Abalone/Protein-folding.html>) (by Molecular Dynamics). .
- [27] Kmiecik, S., and Kolinski, A. (2007). "Characterization of protein-folding pathways by reduced-space modeling" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1941469>). *Proc. Natl. Acad. Sci. U.S.A.* **104** (30): 12330–5. doi:10.1073/pnas.0702265104. PMID 17636132. PMC 1941469.
- [28] <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0009391>
- [29] <http://fold.it/portal/info/science>
- [30] <http://www.stanford.edu/group/pandegroup/folding/about.html>
- [31] <http://boinc.bakerlab.org/rosetta>

Protein dynamics

A **protein domain** is a part of protein sequence and structure that can evolve, function, and exist independently of the rest of the protein chain. Each domain forms a compact three-dimensional structure and often can be independently stable and folded. Many proteins consist of several structural domains. One domain may appear in a variety of evolutionarily related proteins. Domains vary in length from between about 25 amino acids up to 500 amino acids in length. The shortest domains such as zinc fingers are stabilized by metal ions or disulfide bridges. Domains often form functional units, such as the calcium-binding EF hand domain of calmodulin. Because they are self-stable, domains can be "swapped" by genetic engineering between one protein and another to make chimeric proteins.

Background

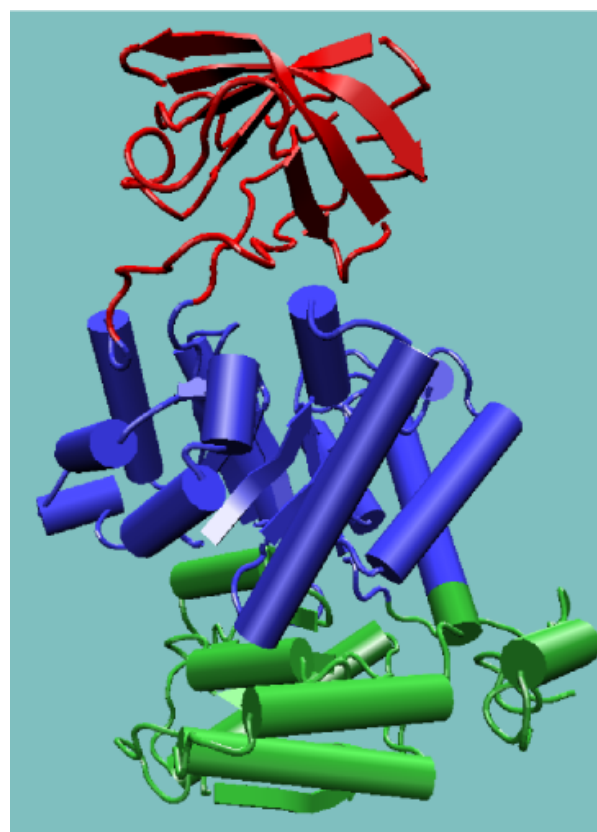
The concept of the **domain** was first proposed in 1973 by Wetlaufer after X-ray crystallographic studies of hen lysozyme ^[2] and papain ^[3] and by limited proteolysis studies of immunoglobulins ^[4] ^[5]. Wetlaufer defined domains as stable units of protein structure that could fold autonomously. In the past domains have been described as units of:

- compact structure^[6]
- function and evolution^[7]
- folding^[8].

Each definition is valid and will often overlap, i.e. a compact structural domain that is found amongst diverse proteins is likely to fold independently within its structural environment. Nature often brings several domains together to form multidomain and multifunctional proteins with a vast number of possibilities ^[9]. In a multidomain protein, each domain may fulfil its own function independently, or in a concerted manner with its neighbours. Domains can either serve as modules for building up large assemblies such as virus particles or muscle fibres, or can provide specific catalytic or binding sites as found in enzymes or regulatory proteins.

An appropriate example is pyruvate kinase, a glycolytic enzyme that plays an important role in regulating the flux from fructose-1,6-biphosphate to pyruvate. It contains an all- β regulatory domain, an α/β -substrate binding domain and an α/β -nucleotide binding domain, connected by several polypeptide linkers ^[10] (see figure, right). Each domain in this protein occurs in diverse sets of protein families.

The central α/β -barrel substrate binding domain is one of the most common enzyme folds. It is seen in many different enzyme families catalysing completely unrelated reactions ^[11]. The α/β -barrel is commonly called the TIM barrel named after triose phosphate isomerase, which was the first such structure to be solved ^[12]. It is currently classified into 26 homologous families in the CATH domain database ^[13]. The TIM barrel is formed from a sequence of β - α - β motifs closed by the first and last strand hydrogen bonding together, forming an eight stranded



Pyruvate kinase, a protein from three domains (PDB 1pkn ^[1])

barrel. There is debate about the evolutionary origin of this domain. One study has suggested that a single ancestral enzyme could have diverged into several families^[14], while another suggests that a stable TIM-barrel structure has evolved through convergent evolution^[15].

The TIM-barrel in pyruvate kinase is 'discontinuous', meaning that more than one segment of the polypeptide is required to form the domain. This is likely to be the result of the insertion of one domain into another during the protein's evolution. It has been shown from known structures that about a quarter of structural domains are discontinuous.^{[16] [17]} The inserted β -barrel regulatory domain is 'continuous', made up of a single stretch of polypeptide.

Covalent association of two domains represents a functional and structural advantage since there is an increase in stability when compared with the same structures non-covalently associated^[18]. Other, advantages are the protection of intermediates within inter-domain enzymatic clefts that may otherwise be unstable in aqueous environments, and a fixed stoichiometric ratio of the enzymatic activity necessary for a sequential set of reactions^[19].

Domains are units of protein structure

Primary structure

The primary structure (string of amino acids) of a protein encodes its uniquely folded 3D conformation.^[20] The most important factor governing the folding of a protein into 3D structure is the distribution of polar and non-polar side chains.^[21] Folding is driven by the burial of hydrophobic side chains into the interior of the molecule so to avoid contact with the aqueous environment.

Sequence alignment is an important tool for determining domains.

Secondary structure

Generally proteins have a core of hydrophobic residues surrounded by a shell of hydrophilic residues. Since the peptide bonds themselves are polar they are neutralised by hydrogen bonding with each other when in the hydrophobic environment. This gives rise to regions of the polypeptide that form regular 3D structural patterns called 'secondary structure'. There are two main types of secondary structure:

- α -helices
- β -sheet

Secondary structure motifs

Some simple combinations of secondary structure elements have been found to frequently occur in protein structure and are referred to as 'super-secondary structure' or motifs. For example, the β -hairpin motif consists of two adjacent antiparallel β -strands joined by a small loop. It is present in most antiparallel β structures both as an isolated ribbon and as part of more complex β -sheets. Another common super-secondary structure is the β - α - β motif, which is frequently used to connect two parallel β -strands. The central α -helix connects the C-termini of the first strand to the N-termini of the second strand, packing its side chains against the β -sheet and therefore shielding the hydrophobic residues of the β -strands from the surface.

Tertiary structure

Several motifs pack together to form compact, local, semi-independent units called domains.^[6] The overall 3D structure of the polypeptide chain is referred to as the protein's 'tertiary structure'. Domains are the fundamental units of tertiary structure, each domain containing an individual hydrophobic core built from secondary structural units connected by loop regions. The packing of the polypeptide is usually much tighter in the interior than the exterior of the domain producing a solid-like core and a fluid-like surface.^[22] In fact, core residues are often conserved in a protein family, whereas the residues in loops are less conserved, unless they are involved in the protein's function. Protein tertiary structure can be divided into four main classes based on the secondary structural content of the domain.^[23]

- All- α domains have a domain core built exclusively from α -helices. This class is dominated by small folds, many of which form a simple bundle with helices running up and down.
- All- β domains have a core comprising of antiparallel β -sheets, usually two sheets packed against each other. Various patterns can be identified in the arrangement of the strands, often giving rise to the identification of recurring motifs, for example the Greek key motif.^[24]
- α + β domains are a mixture of all- α and all- β motifs. Classification of proteins into this class is difficult because of overlaps to the other three classes and therefore is not used in the CATH domain database.^[13]
- α / β domains are made from a combination of β - α - β motifs that predominantly form a parallel β -sheet surrounded by amphipathic α -helices. The secondary structures are arranged in layers or barrels.

Structural alignment is an important tool for determining domains.

Domains have limits on size

Domains have limits on size.^[25] The size of individual structural domains varies from 36 residues in E-selectin to 692 residues in lipoxxygenase-1,^[16] but the majority, 90%, have less than 200 residues^[26] with an average of approximately 100 residues.^[27] Very short domains, less than 40 residues, are often stabilised by metal ions or disulfide bonds. Larger domains, greater than 300 residues, are likely to consist of multiple hydrophobic cores.^[28]

Relationship between primary and tertiary structure

Modules

Nature is a tinkerer and not an inventor,^[29] new sequences are adapted from pre-existing sequences rather than invented. Domains are the common material used by nature to generate new sequences, they can be thought of as genetically mobile units, referred to as 'modules'. Often, the C and N termini of domains are close together in space, allowing them to easily be "slotted into" parent structures during the process of evolution. Many domain families are found in all three forms of life, Archaea, Bacteria and Eukarya. Domains that are repeatedly found in diverse proteins are often referred to as modules, examples can be found among extracellular proteins associated with clotting, fibrinolysis, complement, the extracellular matrix, cell surface adhesion molecules and cytokine receptors.^[30]

Protein families

Molecular evolution gives rise to families of related proteins with similar sequence and structure. However, sequence similarities can be extremely low between proteins that share the same structure. Protein structures may be similar because proteins have diverged from a common ancestor. Alternatively, some folds may be more favored than others as they represent stable arrangements of secondary structures and some proteins may converge towards these folds over the course of evolution. There are currently about 45,000 experimentally determined protein 3D structures deposited within the Protein Data Bank (PDB).^[31] However this set contains a lot of identical or very similar structures. All proteins should be classified to structural families to understand their evolutionary relationships.

Structural comparisons are best achieved at the domain level. For this reason many algorithms have been developed to automatically assign domains in proteins with known 3D structure, see 'Domain definition from structural co-ordinates'.

Super-folds

The CATH domain database classifies domains into approximately 800 fold families, ten of these folds are highly populated and are referred to as 'super-folds'. Super-folds are defined as folds for which there are at least three structures without significant sequence similarity.^[32] The most populated is the α/β -barrel super-fold as described previously.

Multidomain proteins

The majority of genomic proteins, two-thirds in unicellular organisms and more than 80% in metazoa, are multidomain proteins created as a result of gene duplication events.^[33] Many domains in multidomain structures could have once existed as independent proteins. More and more domains in eukaryotic multidomain proteins can be found as independent proteins in prokaryotes.^[34] For example, vertebrates have a multi-enzyme polypeptide containing the GAR synthetase, AIR synthetase and GAR transformylase modules (GARs-AIRs-GARt; GAR: glycylamide ribonucleotide synthetase/transferase; AIR: aminoimidazole ribonucleotide synthetase). In insects, the polypeptide appears as GARs-(AIRs)₂-GARt, in yeast GARs-AIRs is encoded separately from GARt, and in bacteria each domain is encoded separately.^[35]

Origin

Multidomain proteins are likely to have emerged from a selective pressure during evolution to create new functions. Various proteins have diverged from common ancestors by different combinations and associations of domains. Modular units frequently move about, within and between biological systems through mechanisms of genetic shuffling:

- transposition of mobile elements including horizontal transfers (between species),^[36]
- gross rearrangements such as inversions, translocations, deletions and duplications;
- homologous recombination;
- slippage of DNA polymerase during replication.

Difference in proliferation

It is likely that all these and organisms. For example, the ABC transporter domain constitutes one of the largest domain families that appear in all organisms.^[35] Many other families that appear in all organisms show much less proliferation. These include metabolic enzymes and components of translational apparatus.

Types of organisation

The simplest multidomain organisation seen in proteins is that of a single domain repeated in tandem.^[37] The domains may interact with each other or remain isolated, like beads on string. The giant 30,000 residue muscle protein titin comprises about 120 fibronectin-III-type and Ig-type domains.^[38] In the serine proteases, a gene duplication event has led to the formation of a two β -barrel domain enzyme.^[39] The repeats have diverged so widely that there is no obvious sequence similarity between them. The active site is located at a cleft between the two β -barrel domains, in which functionally important residues are contributed from each domain. Genetically engineered mutants of the chymotrypsin serine protease were shown to have some proteinase activity even though their active site residues were abolished and it has therefore been postulated that the duplication event enhanced the enzyme's activity.^[39]

Connectivity

Modules frequently display different connectivity relationships, as illustrated by the kinesins and ABC transporters. The kinesin motor domain can be at either end of a polypeptide chain that includes a coiled-coil region and a cargo domain.^[40] ABC transporters are built with up to four domains consisting of two unrelated modules, ATP-binding cassette and an integral membrane module, arranged in various combinations.

Domain insertion

Not only do domains recombine, but there are many examples of a domain having been inserted into another. Sequence or structural similarities to other domains demonstrate that homologues of inserted and parent domains can exist independently. An example is that of the 'fingers' inserted into the 'palm' domain within the polymerases of the Pol I family.^[41]

Difference between structural and evolutionary domain

Since a domain can be inserted into another, there should always be at least one continuous domain in a multidomain protein. This is the main difference between definitions of structural domains and evolutionary/functional domains. An evolutionary domain will be limited to one or two connections between domains, whereas structural domains can have unlimited connections, within a given criterion of the existence of a common core. Several structural domains could be assigned to an evolutionary domain.

Domains are autonomous folding units

Folding

History

Protein folding - the unsolved problem

Since the seminal work of Anfinsen over forty years ago,^[20] the goal to completely understand the mechanism by which a polypeptide rapidly folds into its stable native conformation remains elusive. Many experimental folding studies have contributed much to our understanding, but the principles that govern protein folding are still based on those discovered in the very first studies of folding. Anfinsen showed that the native state of a protein is thermodynamically stable, the conformation being at a global minimum of its free energy.

Folding pathway

Folding is a directed search of conformational space allowing the protein to fold on a biologically feasible time scale. The Levinthal paradox states that if an averaged sized protein would sample all possible conformations before finding the one with the lowest energy, the whole process would take billions of years.^[42] Proteins typically fold within 0.1 and 1000 seconds, therefore the protein folding process must be directed some way through a specific folding pathway. The forces that direct this search are likely to be a combination of local and global influences whose effects are felt at various stages of the reaction.^[43]

Advances in experimental and theoretical studies have shown that folding can be viewed in terms of energy landscapes,^{[44] [45]} where folding kinetics is considered as a progressive organisation of an ensemble of partially folded structures through which a protein passes on its way to the folded structure. This has been described in terms of a folding funnel, in which an unfolded protein has a large number of conformational states available and there are fewer states available to the folded protein. A funnel implies that for protein folding there is a decrease in energy and loss of entropy with increasing tertiary structure formation. The local roughness of the funnel reflects kinetic traps, corresponding to the accumulation of misfolded intermediates. A folding chain progresses toward lower intra-chain free-energies by increasing its compactness. The chains conformational options become increasingly narrowed

ultimately toward one native structure.

Advantage of domains in protein folding

The organisation of large proteins by structural domains represents an advantage for protein folding, with each domain being able to individually fold, accelerating the folding process and reducing a potentially large combination of residue interactions. Furthermore, given the observed random distribution of hydrophobic residues in proteins,^[46] domain formation appears to be the optimal solution for a large protein to bury its hydrophobic residues while keeping the hydrophilic residues at the surface.^{[47] [48]}

However, the role of inter-domain interactions in protein folding and in energetics of stabilisation of the native structure, probably differs for each protein. In T4 lysozyme, the influence of one domain on the other is so strong that the entire molecule is resistant to proteolytic cleavage. In this case, folding is a sequential process where the C-terminal domain is required to fold independently in an early step, and the other domain requires the presence of the folded C-terminal domain for folding and stabilisation.^[49]

It has been found that the folding of an isolated domain can take place at the same rate or sometimes faster than that of the integrated domain.^[50] Suggesting that unfavourable interactions with the rest of the protein can occur during folding. Several arguments suggest that the slowest step in the folding of large proteins is the pairing of the folded domains.^[28] This is either because the domains are not folded entirely correctly or because the small adjustments required for their interaction are energetically unfavourable,^[51] such as the removal of water from the domain interface.

Domains and quaternary structure

About quaternary structures

Many proteins have a quaternary structure, which consists of several polypeptide chains that associate into an oligomeric molecule. Each polypeptide chain in such a protein is called a subunit. Hemoglobin, for example, consists of two α and two β subunits. Each of the four chains has an all- α globin fold with a heme pocket.

Domain swapping

Domain swapping is a mechanism for forming oligomeric assemblies.^[52] In domain swapping, a secondary or tertiary element of a monomeric protein is replaced by the same element of another protein. Domain swapping can range from secondary structure elements to whole structural domains. It also represents a model of evolution for functional adaptation by oligomerisation, e.g. oligomeric enzymes that have their active site at subunit interfaces.^[53]

Domains and protein flexibility

The presence of multiple domains in proteins gives rise to a great deal of flexibility and mobility, leading to **protein domain dynamics**. Domain motions can be inferred by comparing structures of a protein in different environments, or directly observed using spectra^[54] measured by neutron spin echo spectroscopy. One of the largest observed domain motions is the 'swivelling' mechanism in pyruvate phosphate dikinase. The phosphoinositide domain swivels between two states in order to bring a phosphate group from the active site of the nucleotide binding domain to that of the phosphoenolpyruvate/pyruvate domain.^[55] The phosphate group is moved over a distance of 45Å involving a domain motion of about 100 degrees around a single residue. Domain motions are important for:^[56]

- catalysis;
- regulatory activity;
- transport of metabolites;
- formation of protein assemblies; and

- cellular locomotion.

In enzymes, the closure of one domain onto another captures a substrate by an induced fit, allowing the reaction to take place in a controlled way. A detailed analysis by Gerstein led to the classification of two basic types of domain motion; hinge and shear.^[56] Only a relatively small portion of the chain, namely the inter-domain linker and side chains undergo significant conformational changes upon domain rearrangement.^[57]

Hinges by secondary structures

A study by Hayward^[58] found that the termini of α -helices and β -sheets form hinges in a large number of cases. Many hinges were found to involve two secondary structure elements acting like hinges of a door, allowing an opening and closing motion to occur. This can arise when two neighbouring strands within a β -sheet situated in one domain, diverge apart as they join the other domain. The two resulting termini then form the bending regions between the two domains. α -helices that preserve their hydrogen bonding network when bent are found to behave as mechanical hinges, storing 'elastic energy' that drives the closure of domains for rapid capture of a substrate.^[58]

Helical to extended conformation

The interconversion of helical and extended conformations at the site of a domain boundary is not uncommon. In calmodulin, torsion angles change for five residues in the middle of a domain linking α -helix. The helix is split into two, almost perpendicular, smaller helices separated by four residues of an extended strand.^{[59] [60]}

Shear motions

Shear motions involve a small sliding movement of domain interfaces, controlled by the amino acid side chains within the interface. Proteins displaying shear motions often have a layered architecture: stacking of secondary structures. The interdomain linker has merely the role of keeping the domains in close proximity.

Domain definition from structural co-ordinates

The importance of domains as structural building blocks and elements of evolution has brought about many automated methods for their identification and classification in proteins of known structure. Automatic procedures for reliable domain assignment is essential for the generation of the domain databases, especially as the number of protein structures is increasing. Although the boundaries of a domain can be determined by visual inspection, construction of an automated method is not straightforward. Problems occur when faced with domains that are discontinuous or highly associated.^[61] The fact that there is no standard definition of what a domain really is has meant that domain assignments have varied enormously, with each researcher using a unique set of criteria.^[62]

A structural domain is a compact, globular sub-structure with more interactions within it than with the rest of the protein.^[57] Therefore, a structural domain can be determined by two visual characteristics; its compactness and its extent of isolation.^[63] Measures of local compactness in proteins have been used in many of the early methods of domain assignment^{[64] [65] [66] [67]} and in several of the more recent methods.^{[26] [68] [69] [70] [71]}

Considering proteins as small segments

One of the first algorithms^[64] used a C α -C α distance map together with a hierarchical clustering routine that considered proteins as several small segments, 10 residues in length. The initial segments were clustered one after another based on inter-segment distances; segments with the shortest distances were clustered and considered as single segments thereafter. The stepwise clustering finally included the full protein. Go^[67] also exploited the fact that inter-domain distances are normally larger than intra-domain distances; all possible C α -C α distances were represented as diagonal plots in which there were distinct patterns for helices, extended strands and combinations of secondary structures.

Sowdhamini and Blundell's method

The method by Sowdhamini and Blundell clusters secondary structures in a protein based on their C α -C α distances and identifies domains from the pattern in their dendrograms.^[61] As the procedure does not consider the protein as a continuous chain of amino acids there are no problems in treating discontinuous domains. Specific nodes in these dendrograms are identified as tertiary structural clusters of the protein, these include both super-secondary structures and domains. The DOMAK algorithm is used to create the 3Dee domain database.^[69] It calculates a 'split value' from the number of each type of contact when the protein is divided arbitrarily into two parts. This split value is large when the two parts of the structure are distinct.

Method of Wodak and Janin

The method of Wodak and Janin^[72] was based on the calculated interface areas between two chain segments repeatedly cleaved at various residue positions. Interface areas were calculated by comparing surface areas of the cleaved segments with that of the native structure. Potential domain boundaries can be identified at a site where the interface area was at a minimum.

Other methods have used measures of solvent accessibility to calculate compactness.^{[26] [73] [74]}

PUU algorithm

The PUU algorithm^[17] incorporates a harmonic model used to approximate inter-domain dynamics. The underlying physical concept is that many rigid interactions will occur within each domain and loose interactions will occur between domains. This algorithm is used to define domains in the FSSP domain database.^[68]

DETECTIVE

Swindells (1995) developed a method, DETECTIVE, for identification of domains in protein structures based on the idea that domains have a hydrophobic interior. Deficiencies were found to occur when hydrophobic cores from different domains continue through the interface region.

RigidFinder

RigidFinder^[75] is a novel method for identification of protein rigid blocks (domains and loops) from two different conformations. Rigid blocks are defined as blocks where all inter residue distances are conserved across conformations.

PiSQRD

The PiSQRD^[76] web-server allows users to optimally subdivide single-chain or multimeric proteins into domains that behave approximately as rigid units in the course of protein structural fluctuations^{[77] [78]}. The best rigid-body decomposition is found using the lowest-energy collective modes of the system. By default the latter are calculated through an elastic network model, or can be uploaded by the user.

Example domains

Armadillo repeats

named after the β -catenin-like Armadillo protein of the fruit fly *Drosophila*.

Basic Leucine zipper domain (bZIP domain)

is found in many DNA-binding eukaryotic proteins. One part of the domain contains a region that mediates sequence-specific DNA-binding properties and the Leucine zipper that is required for the dimerization of two DNA-binding regions. The DNA-binding region comprises a number of basic aminoacids such as arginine and lysine

Cadherin repeats

Cadherins function as Ca^{2+} -dependent cell-cell adhesion proteins. Cadherin domains are extracellular regions which mediate cell-to-cell homophilic binding between cadherins on the surface of adjacent cells.

Death effector domain (DED)

allows protein-protein binding by homotypic interactions (DED-DED). Caspase proteases trigger apoptosis via proteolytic cascades. Pro-Caspase-8 and pro-caspase-9 bind to specific adaptor molecules via DED domains and this leads to autoactivation of caspases.

EF hand

a helix-turn-helix structural motif found in each structural domain of the signaling protein calmodulin and in the muscle protein troponin-C.

Immunoglobulin-like domains

are found in proteins of the immunoglobulin superfamily (IgSF).^[79] They contain about 70-110 amino acids and are classified into different categories (IgV, IgC1, IgC2 and IgI) according to their size and function. They possess a characteristic fold in which two beta sheets form a "sandwich" that is stabilized by interactions between conserved cysteines and other charged amino acids. They are important for protein-to-protein interactions in processes of cell adhesion, cell activation, and molecular recognition. These domains are commonly found in molecules with roles in the immune system.

Phosphotyrosine-binding domain (PTB)

PTB domains usually bind to phosphorylated tyrosine residues. They are often found in signal transduction proteins. PTB-domain binding specificity is determined by residues to the amino-terminal side of the phosphotyrosine. Examples: the PTB domains of both SHC and IRS-1 bind to a NPXpY sequence. PTB-containing proteins such as SHC and IRS-1 are important for insulin responses of human cells.

Pleckstrin homology domain (PH)

PH domains bind phosphoinositides with high affinity. Specificity for PtdIns(3)P, PtdIns(4)P, PtdIns(3,4)P₂, PtdIns(4,5)P₂, and PtdIns(3,4,5)P₃ have all been observed. Given the fact that phosphoinositides are sequestered to various cell membranes (due to their long lipophilic tail) the PH domains usually causes recruitment of the protein in question to a membrane where the protein can exert a certain function in cell signalling, cytoskeletal reorganization or membrane trafficking.

Src homology 2 domain (SH2)

SH2 domains are often found in signal transduction proteins. SH2 domains confer binding to phosphorylated tyrosine (pTyr). Named after the phosphotyrosine binding domain of the src viral oncogene, which is itself a tyrosine kinase. *See also*: SH3 domain.

Zinc finger DNA binding domain (ZnF_GATA)

ZnF_GATA domain-containing proteins are typically transcription factors that usually bind to the DNA sequence [AT]GATA[AG] of promoters.

The preceding text and figures originate from "Predicting Structural Domains in Proteins" George RA, 2002.

See also

- Amino acid
- Binding domain
- CATH
- Conserved domains
- Motif domain
- Eukaryotic Linear Motif
- Protein
- Protein structure
- Protein structure prediction
- Protein structure prediction software
- Protein family
- Structural biology
- Structural Classification of Proteins (SCOP)

Structural domain databases

- 3Dee ^[80]
- CATH ^[81]
- DALI ^[82]
- SCOP ^[83]
- Pawson Lab - Protein interaction domains ^[84]
- Nash Lab - Protein interaction domains in Signal Transduction ^[85]
- Definition and assignment of structural domains in proteins ^[86].

Sequence domain databases

- InterPro ^[87]
 - Pfam ^[88]
 - PROSITE ^[89]
 - ProDom ^[90]
 - SMART ^[91]
 - NCBI Conserved Domain Database ^[92]
 - SUPERFAMILY ^[93] Library of HMMs representing superfamilies and database of (superfamily and family) annotations for all completely sequenced organisms
-

External links

- The Protein Families (Pfam) database clan browser ^[94] provides easy access to information about protein structural domains. A clan contains two or more **Pfam** families that have arisen from a single evolutionary origin.

Key papers

- Bastian, H. C. (1872). *The beginnings of life: being some account of the nature, modes of origin and transformation of lower organisms*. Macmillan and Co., England.
- Berman HM et al. (2000). "The Protein Data Bank" ^[95]. *Nucleic Acids Res* **28** (1): 235–42. doi:10.1093/nar/28.1.235. PMID 10592235. PMC 102472.
- Branden, C.-I. and Tooze, J. (1991). *Introduction to protein structure*. Garland, New York.
- Chothia C. (1992). "Proteins. One thousand families for the molecular biologist". *Nature* **357** (6379): 543–4. doi:10.1038/357543a0. PMID 1608464.
- Das S, Smith TF. (2000). "Identifying nature's protein Lego set". *Adv Protein Chem* **54**: 159–83. PMID 10829228.
- Dietmann S, Park J, Notredame C, Heger A, Lappe M, Holm L. (2001). "A fully automatic evolutionary classification of protein folds: Dali Domain Dictionary version 3" ^[96]. *Nucleic Acids Res* **29** (1): 55–7. doi:10.1093/nar/29.1.55. PMID 11125048. PMC 29815.
- Dill, Ken A.; Chan, Hue Sun (1997). "From Levinthal to pathways to funnels". *Nat Struc Biol* **4** (1): 10. doi:10.1038/nsb0197-10.
- Dyson HJ, Sayre JR, Merutka G, Shin HC, Lerner RA, Wright PE. (1992). "Folding of peptide fragments comprising the complete sequence of proteins. Models for initiation of protein folding. II. Plastocyanin". *J Mol Biol* **226** (3): 819–35. doi:10.1016/0022-2836(92)90634-V. PMID 1507228.
- Fersht AR. (1997). "Nucleation mechanisms in protein folding". *Curr Opin Struct Biol* **7** (1): 3–9. doi:10.1016/S0959-440X(97)80002-4. PMID 9032066.
- George DG, Hunt LT, Barker WC. (1996). "PIR-International Protein Sequence Database". *Methods Enzymol* **266**: 41–59. PMID 8743676.
- George, R. A. (2002) "Predicting Structural Domains in Proteins". Thesis, University College London
- Go M. (1981). "Correlation of DNA exonic regions with protein structural units in haemoglobin". *Nature* **291** (5810): 90–2. doi:10.1038/291090a0. PMID 7231530.
- Hadley, C and Jones, D.T. (1999). "A systematic comparison of protein structure classifications: SCOP, CATH and FSSP". *Struct Fold Des* **7** (9): 1099. doi:10.1016/S0969-2126(99)80177-4.
- Hayward S. (1999). "Structural principles governing domain motions in proteins". *Proteins* **36** (4): 425–35. doi:10.1002/(SICI)1097-0134(19990901)36:4<425::AID-PROT6>3.0.CO;2-S (inactive 2010-03-18). PMID 10450084.
- Heringa J, Argos P. (1991). "Side-chain clusters in protein structures and their role in protein folding". *J Mol Biol* **220** (1): 151–71. doi:10.1016/0022-2836(91)90388-M. PMID 2067014.
- Honig B. (1999). "Protein folding: from the levinthal paradox to structure prediction". *J Mol Biol* **293** (2): 283–93. doi:10.1006/jmbi.1999.3006. PMID 10550209.
- Kim PS, Baldwin RL. (1990). "Intermediates in the folding reactions of small proteins". *Annu Rev Biochem* **59** (1): 631–60. doi:10.1146/annurev.bi.59.070190.003215. PMID 2197986.
- Larsen TM, Laughlin LT, Holden HM, Rayment I, Reed GH. (1994). "Structure of rabbit muscle pyruvate kinase complexed with Mn²⁺, K⁺, and pyruvate". *Biochemistry* **33** (20): 6301–9. doi:10.1021/bi00186a033. PMID 8193145.
- Murvai J, Vlahovicek K, Barta E, Cataletto B, Pongor S. (2000). "The SBASE protein domain library, release 7.0: a collection of annotated protein sequence segments" ^[97]. *Nucleic Acids Res* **28** (1): 260–2. doi:10.1093/nar/28.1.260. PMID 10592241. PMC 102474.

- Murzin AG, Brenner SE, Hubbard T, Chothia C. (1995). "SCOP: a structural classification of proteins database for the investigation of sequences and structures". *J Mol Biol* **247** (4): 536–40. doi:10.1016/S0022-2836(05)80134-2. PMID 7723011.
- Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. (2000). "The structural basis of ribosome activity in peptide bond synthesis". *Science* **289** (5481): 920–30. doi:10.1126/science.289.5481.920. PMID 10937990.
- Janin J, Chothia C. (1985). "Domains in proteins: definitions, location, and structural principles". *Methods Enzymol* **115**: 420–30. PMID 4079796.
- Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. (2000). "SMART: a web-based tool for the study of genetically mobile domains" ^[98]. *Nucleic Acids Res* **28** (1): 231–4. doi:10.1093/nar/28.1.231. PMID 10592234. PMC 102444.
- Siddiqui AS, Dengler U, Barton GJ. (2001). "3Dee: a database of protein structural domains". *Bioinformatics* **17** (2): 200–1. doi:10.1093/bioinformatics/17.2.200. PMID 11238081.
- Srinivasarao GY, Yeh LS, Marzec CR, Orcutt BC, Barker WC, Pfeiffer F. (1999). "Database of protein sequence alignments: PIR-ALN" ^[99]. *Nucleic Acids Res* **27** (1): 284–5. doi:10.1093/nar/27.1.284. PMID 9847202. PMC 148157.
- Tatusov RL et al. (2001). "The COG database: new developments in phylogenetic classification of proteins from complete genomes" ^[100]. *Nucleic Acids Res* **29** (1): 22–8. doi:10.1093/nar/29.1.22. PMID 11125040. PMC 29819.
- Taylor WR, Orengo CA. (1989). "Protein structure alignment". *J Mol Biol* **208** (1): 1–22. doi:10.1016/0022-2836(89)90084-3. PMID 2769748.
- Yang AS, Honig B. (1995). "Free energy determinants of secondary structure formation: I. alpha-Helices". *J Mol Biol* **252** (3): 351–65. doi:10.1006/jmbi.1995.0502. PMID 7563056.
- Yang AS, Honig B. (1995). "Free energy determinants of secondary structure formation: II. Antiparallel beta-sheets". *J Mol Biol* **252** (3): 366–76. doi:10.1006/jmbi.1995.0503. PMID 7563057. }

References

- [1] <http://www.rcsb.org/pdb/explore/explore.do?structureId=1pkn>
- [2] Phillips DC. (1966). "The three-dimensional structure of an enzyme molecule". *Scientific American* **215** (5): 78–90. doi:10.1038/255609a0. PMID 5978599.
- [3] Drenth J, Jansonius JN, Koekoek R, Swen HM, Wolthers BG. (1968). "Structure of papain". *Nature* **218** (5145): 929–32. doi:10.1038/218929a0. PMID 5681232.
- [4] Porter RR. (1973). "Structural studies of immunoglobulins" (<http://www.sciencemag.org/cgi/pmidlookup?view=long&pmid=4122075>). *Science* **180** (87): 713–6. doi:10.1126/science.180.4087.713. PMID 4122075. .
- [5] Edelman GM. (1973). "Antibody structure and molecular immunology". *Science* **180** (88): 830–40. doi:10.1126/science.180.4088.830. PMID 4540988.
- [6] Richardson J. S. (1981). "The anatomy and taxonomy of protein structure" (<http://kinemage.biochem.duke.edu/teaching/anatax/index.html>). *Adv Protein Chem* **34**: 167–339. doi:10.1038/255609a0. PMID 7020376. .
- [7] Bork P. (1991). "Shuffled domains in extracellular proteins". *FEBS Lett* **286** (1-2): 47–54. doi:10.1016/0014-5793(91)80937-X. PMID 1864378.
- [8] Wetlaufer DB. (1973). "Nucleation, rapid folding, and globular intrachain regions in proteins" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=433338>). *Proc Natl Acad Sci USA* **70** (3): 697–701. doi:10.1073/pnas.70.3.697. PMID 4351801. PMC 433338.
- [9] Chothia C. (1992). "Proteins. One thousand families for the molecular biologist". *Nature* **357** (6379): 543–4. doi:10.1038/357543a0. PMID 1608464.
- [10] George RA, Heringa J. (2002). "An analysis of protein domain linkers: their classification and role in protein folding". *Protein Eng* **15** (11): 871–9. doi:10.1093/protein/15.11.871. PMID 12538906.
- [11] Hegyi H, and Gerstein M. (1999). "The relationship between protein structure and function: a comprehensive survey with application to the yeast genome". *J Mol Biol* **288** (1): 147–64. doi:10.1006/jmbi.1999.2661. PMID 10329133.
- [12] Banner et al.; Bloomer, AC; Petsko, GA; Phillips, DC; Pogson, CI; Wilson, IA; Corran, PH; Furth, AJ et al. (1975). ""Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data"". *Nature* **255** (5510): 609–614. doi:10.1038/255609a0. PMID 1134550.

- [13] Orengo CA, Michie AD, Jones S, Jones DT, Swindells MB, Thornton JM. (1997). "CATH--a hierarchic classification of protein domain structures". *Structure* **5** (8): 1093–108. doi:10.1016/S0969-2126(97)00260-8. PMID 9309224.
- [14] Copley, R. R. and Bork, P (2000). ""Homology among (betaalpha)(8) barrels: implications for the evolution of metabolic pathways"". *J Mol Biol* **303** (4): 627–641. doi:10.1006/jmbi.2000.4152. PMID 11054297.
- [15] Lesk AM, Brändén CI, Chothia C. (1989). "Structural principles of alpha/beta barrel proteins: the packing of the interior of the sheet". *Proteins* **5** (2): 139–48. doi:10.1002/prot.340050208. PMID 2664768.
- [16] Jones S, Stewart M, Michie A, Swindells MB, Orengo C, Thornton JM. (1998). "Domain assignment for protein structures using a consensus approach: characterization and analysis" (<http://www.proteinscience.org/cgi/pmidlookup?view=long&pmid=9521098>). *Protein Sci* **7** (2): 233–42. doi:10.1002/pro.5560070202 (inactive 2010-03-18). PMID 9521098. PMC 2143930. .
- [17] Holm L, Sander C. (1994). "Parser for protein folding units". *Proteins* **19** (3): 256–68. doi:10.1002/prot.340190309. PMID 7937738.
- [18] Ghélis C, Yon JM. (1979). "[Conformational coupling between structural units. A decisive step in the functional structure formation]". *C R Seances Acad Sci D* **289** (2): 197–9. PMID 117925.
- [19] Ostermeier M, Benkovic SJ. (2000). "Evolution of protein function by domain swapping". *Adv Protein Chem* **55**: 29–77. PMID 11050932.
- [20] ANFINSEN CB, HABER E, SELA M, WHITE FH Jr. (1961). "The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=223141>). *Proc Natl Acad Sci USA* **47** (9): 1309–14. doi:10.1073/pnas.47.9.1309. PMID 13683522. PMC 223141. 10.1073/pnas.47.9.1309}}
- [21] Cordes, M. H., Davidson, A. R., and Sauer, R. T (1996). ""Sequence space, folding and protein design"". *Curr Opin Struct Biol* **6** (1): 3–10. doi:10.1016/S0959-440X(96)80088-1. PMID 8696970.
- [22] Zhou, Y., Vitkup, D., and Karplus, M (1999). ""Native proteins are surface-molten solids: application of the Lindemann criterion for the solid versus liquid state"". *J Mol Biol* **285** (4): 1371–1375. doi:10.1006/jmbi.1998.2374. PMID 9917381.
- [23] Levitt M, Chothia C. (1976). "Structural patterns in globular proteins". *Nature* **261** (5561): 552–8. doi:10.1038/261552a0. PMID 934293.
- [24] Hutchinson EG, Thornton JM. (1993). "The Greek key motif: extraction, classification and analysis". *Protein Eng* **6** (3): 233–45. doi:10.1093/protein/6.3.233. PMID 8506258.
- [25] Savageau MA. (1986). "Proteins of Escherichia coli come in sizes that are multiples of 14 kDa: domain concepts and evolutionary implications" (<http://www.pnas.org/cgi/pmidlookup?view=long&pmid=3513170>). *Proc Natl Acad Sci USA* **83** (5): 1198–202. doi:10.1073/pnas.83.5.1198. PMID 3513170. PMC 323042. .
- [26] Islam SA, Luo J, Sternberg MJ. (1995). "Identification and analysis of domains in proteins". *Protein Eng* **8** (6): 513–25. doi:10.1093/protein/8.6.513. PMID 8532675.
- [27] Wheelan, S. J. and Marchler-Bauer, A. and Bryant, S. H. (2000). "Domain size distributions can predict domain boundaries". *Bioinformatics* **16** (7): 613–618. doi:10.1093/bioinformatics/16.7.613. PMID 11038331.
- [28] Garel, J. (1992). "Folding of large proteins: Multidomain and multisubunit proteins". In Creighton, T., editor, *Protein Folding*, pages 405–454. W.H. Freeman and Company, New York, first edition.
- [29] Jacob F. (1977). "Evolution and tinkering". *Science* **196** (4295): 1161–6. doi:10.1126/science.860134. PMID 860134.
- [30] Campbell ID, Downing AK. (1994). "Building protein structure and function from modular units". *Trends Biotechnol* **12** (5): 168–72. doi:10.1016/0167-7799(94)90078-7. PMID 7764899.
- [31] <http://www.pdb.org/>
- [32] Orengo CA, Jones DT, Thornton JM. (1994). "Protein superfamilies and domain superfolds". *Nature* **372** (6507): 631–4. doi:10.1038/372631a0. PMID 7990952.
- [33] Apic, G., Gough, J., and Teichmann, S. A (2001). ""Domain combinations in archaeal, eubacterial and eukaryotic proteomes"". *J Mol Biol* **310** (2): 311–325. doi:10.1006/jmbi.2001.4776. PMID 11428892.
- [34] Davidson JN, Chen KC, Jamison RS, Musmanno LA, Kern CB. (1993). "The evolutionary history of the first three enzymes in pyrimidine biosynthesis". *Bioessays* **15** (3): 157–64. doi:10.1002/bies.950150303. PMID 8098212.
- [35] Henikoff S, Greene EA, Pietrovski S, Bork P, Attwood TK, Hood L. (1997). "Gene families: the taxonomy of protein paralogs and chimeras". *Science* **278** (5338): 609–14. doi:10.1126/science.278.5338.609. PMID 9381171.
- [36] Bork, P. and Doolittle, R. F (1992). ""Proposed acquisition of an animal protein domain by bacteria2" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=50050>). *Proc Natl Acad Sci USA* **89** (19): 8990–8994. doi:10.1073/pnas.89.19.8990. PMID 1409594. PMC 50050.
- [37] Heringa J. (1998). "Detection of internal repeats: how common are they?". *Curr Opin Struct Biol* **8** (3): 338–45. doi:10.1016/S0959-440X(98)80068-7. PMID 9666330.
- [38] Politou, A. S., Gautel, M., Improta, S., Vangelista, L., and Pastore, A (1996). ""The elastic I-band region of titin is assembled in a 'modular' fashion by weakly interacting Ig-like domains"". *J Mol Biol* **255** (4): 604–616. doi:10.1006/jmbi.1996.0050. PMID 8568900.
- [39] McLachlan, A. D (1979). ""Gene duplications in the structural evolution of chymotrypsin"". *J Mol Biol* **128** (1): 49–79. doi:10.1016/0022-2836(79)90308-5. PMID 430571.
- [40] Moore JD, Endow SA. (1996). "Kinesin proteins: a phylum of motors for microtubule-based motility". *Bioessays* **18** (3): 207–19. doi:10.1002/bies.950180308. PMID 8867735.
- [41] Russell, R. B (1994). ""Domain insertion"". *Protein Eng* **7** (12): 1407–1410. doi:10.1093/protein/7.12.1407. PMID 7716150.
- [42] Levinthal, C. (1968). "Are there pathways for protein folding?" (<http://www.biochem.wisc.edu/courses/biochem704/Reading/Levinthal1968.pdf>). *J Chim Phys* **65**: 44–45. .

- [43] Dill KA. (1999). "Polymer principles and protein folding" (<http://www.proteinscience.org/cgi/pmidlookup?view=long&pmid=10386867>). *Protein Sci* **8** (6): 1166–80. doi:10.1110/ps.8.6.1166. PMID 10386867. PMC 2144345. .
- [44] Leopold PE, Montal M, Onuchic JN. (1992). "Protein folding funnels: a kinetic approach to the sequence-structure relationship" (<http://www.pnas.org/cgi/pmidlookup?view=long&pmid=1528885>). *Proc Natl Acad Sci USA* **89** (18): 8721–5. doi:10.1073/pnas.89.18.8721. PMID 1528885. PMC 49992. .
- [45] Dill KA, Chan HS. (1997). "From Levinthal to pathways to funnels". *Nat Struct Biol* **4** (1): 10–9. doi:10.1038/nsb0197-10. PMID 8989315.
- [46] White SH, Jacobs RE. (1990). "Statistical distribution of hydrophobic residues along the length of protein chains. Implications for protein folding and evolution" (<http://www.biophysj.org/cgi/pmidlookup?view=long&pmid=2188687>). *Biophys J* **57** (4): 911–21. doi:10.1016/S0006-3495(90)82611-4. PMID 2188687. PMC 1280792. .
- [47] George RA, Heringa J. (2002). "SnapDRAGON: a method to delineate protein structural domains from sequence data". *J Mol Biol* **316** (3): 839–51. doi:10.1006/jmbi.2001.5387. PMID 11866536.
- [48] George RA, Lin K, Heringa J. (2005). "Scooby-domain: prediction of globular domains in protein sequence" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=15980446>). *Nucleic Acids Res* **33** (Web Server issue): W160–3. doi:10.1093/nar/gki381. PMID 15980446. PMC 1160142. .
- [49] Desmadril, M. and Yon, J. M (1981). ""Existence of intermediates in the refolding of T4 lysozyme at pH 7.4"". *Biochem Biophys Res Commun* **101** (2): 563–569. doi:10.1016/0006-291X(81)91296-1. PMID 7306096.
- [50] Teale JM, Benjamin DC. (1977). "Antibody as immunological probe for studying refolding of bovine serum albumin. Refolding within each domain". *J Biol Chem* **252** (13): 4521–6. PMID 873903.
- [51] Creighton, T. E. (1983). *Proteins: Structures and molecular properties*. Freeman, New York. Second edition.
- [52] Bennett MJ, Schlunegger MP, Eisenberg D. (1995). "3D domain swapping: a mechanism for oligomer assembly" (<http://www.proteinscience.org/cgi/pmidlookup?view=long&pmid=8580836>). *Protein Sci* **4** (12): 2455–68. doi:10.1002/pro.5560041202. PMID 8580836. PMC 2143041. .
- [53] Heringa J, Taylor WR. (1997). "Three-dimensional domain duplication, swapping and stealing". *Curr Opin Struct Biol* **7** (3): 416–21. doi:10.1016/S0959-440X(97)80060-7. PMID 9204285.
- [54] Bu Z, Biehl R, Monkenbusch M, Richter D, D.J.E. Callaway (2005). "Coupled protein domain motion in Taq polymerase revealed by neutron spin-echo spectroscopy." (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1345721>). *Proc Natl Acad Sci USA* **102** (49): 17646–17651. doi:10.1073/pnas.0503388102. PMID 16306270. PMC 1345721.
- [55] Herzberg O et al. (1996). "Swiveling-domain mechanism for enzymatic phosphotransfer between remote reaction sites" (<http://www.pnas.org/cgi/pmidlookup?view=long&pmid=8610096>). *Proc Natl Acad Sci USA* **93** (7): 2652–7. doi:10.1073/pnas.93.7.2652. PMID 8610096. PMC 39685. .
- [56] Gerstein M, Lesk AM, Chothia C. (1994). "Structural mechanisms for domain movements in proteins". *Biochemistry* **33** (22): 6739–49. doi:10.1021/bi00188a001. PMID 8204609.
- [57] Janin, J. and Wodak, S. J (1983). "Structural domains in proteins and their role in the dynamics of protein function". *Prog Biophys Mol Biol* **42** (1): 21–78. doi:10.1016/0079-6107(83)90003-2. PMID 6353481.
- [58] Hayward, 1999
- [59] Meador, W. E., Means, A. R., and Quirocho, F. A (1992). ""Target enzyme recognition by calmodulin: 2.4A structure of a calmodulin-peptide complex"". *Science* **257** (5074): 1251–1255. doi:10.1126/science.1519061. PMID 1519061.
- [60] Ikura, M., Clore, G. M., Gronenborn, A. M., Zhu, G., Klee, C. B., and Bax, A (1992). ""Solution structure of a calmodulin-target peptide complex by multidimensional NMR"". *Science* **256** (5057): 632–638. doi:10.1126/science.1585175. PMID 1585175.
- [61] Sowdhamini R, Blundell TL. (1995). "An automatic method involving cluster analysis of secondary structures for the identification of domains in proteins" (<http://www.proteinscience.org/cgi/pmidlookup?view=long&pmid=7795532>). *Protein Sci* **4** (3): 506–20. doi:10.1002/pro.5560040317 (inactive 2010-03-18). PMID 7795532. PMC 2143076. .
- [62] Swindells, M. B (1995). ""A procedure for detecting structural domains in proteins"" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2142966>). *Protein Sci* **4** (1): 103–112. doi:10.1002/pro.5560040113 (inactive 2010-03-18). PMID 7773168. PMC 2142966.
- [63] Tsai CJ, Nussinov R. (1997). "Hydrophobic folding units derived from dissimilar monomer structures and their interactions" (<http://www.proteinscience.org/cgi/pmidlookup?view=long&pmid=9007974>). *Protein Sci* **6** (1): 24–42. doi:10.1002/pro.5560060104 (inactive 2010-03-18). PMID 9007974. PMC 2143523. .
- [64] Crippen, G. M (1978). ""The tree structural organisation of proteins"". *J Mol Biol* **126** (3): 315–332. doi:10.1016/0022-2836(78)90043-8. PMID 745231.
- [65] Rossmann MG, Moras D, Olsen KW. (1974). "Chemical and biological evolution of nucleotide-binding protein". *Nature* **250** (463): 194–9. doi:10.1038/250194a0. PMID 4368490.
- [66] Rose GD. (1979). "Hierarchic organization of domains in globular proteins". *J Mol Biol* **134** (3): 447–70. doi:10.1016/0022-2836(79)90363-2. PMID 537072.
- [67] Go N, Taketomi H. (1978). "Respective roles of short- and long-range interactions in protein folding" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=411294>). *Proc Natl Acad Sci USA* **75** (2): 559–63. doi:10.1073/pnas.75.2.559. PMID 273218. PMC 411294.
- [68] Holm L, Sander C. (1997). "Dali/FSSP classification of three-dimensional protein folds" (<http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=9016542>). *Nucleic Acids Res* **25** (1): 231–4. doi:10.1093/nar/25.1.231. PMID 9016542. PMC 146389. .

- [69] Siddiqui AS, Barton GJ. (1995). "Continuous and discontinuous domains: an algorithm for the automatic generation of reliable protein domain definitions" (<http://www.proteinscience.org/cgi/pmidlookup?view=long&pmid=7663343>). *Protein Sci* **4** (5): 872–84. doi:10.1002/pro.5560040507 (inactive 2010-03-18). PMID 7663343. PMC 2143117. .
- [70] Zehfus, M. H (1997). ""Identification of compact, hydrophobically stabilized domains and modules containing multiple peptide chains"" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2143719>). *Protein Sci* **6** (6): 1210–1219. doi:10.1002/pro.5560060609. PMID 9194181. PMC 2143719.
- [71] Taylor WR. (1999). "Protein structural domain identification". *Protein Eng* **12** (3): 203–16. doi:10.1093/protein/12.3.203. PMID 10235621.
- [72] Wodak, S. J. and Janin, J (1981). ""Location of structural domains in protein"". *Biochemistry* **20** (23): 6544–6552. doi:10.1021/bi00526a005. PMID 7306523.
- [73] Rashin, 1985
- [74] Zehfus MH, Rose GD. (1986). "Compact units in proteins". *Biochemistry* **25** (19): 5759–65. doi:10.1021/bi00367a062. PMID 3778881.
- [75] <http://rigidfinder.molmovdb.org>
- [76] <http://pisqrd.escience-lab.org>
- [77] Potestio, R., Pontiggia, F. and Micheletti, C. (2009). "Coarse-grained description of protein internal dynamics: an optimal strategy for decomposing proteins in rigid subunits." ([http://www.cell.com/biophysj/abstract/S0006-3495\(09\)00781-4](http://www.cell.com/biophysj/abstract/S0006-3495(09)00781-4)). *Biophysical Journal* **96** (12): 4993–5002. doi:10.1016/j.bpj.2009.03.051. PMID 19527659. PMC 2712024. .
- [78] Aleksiev, T., Potestio, R., Pontiggia, F., Cozzini, S. and Micheletti, C. (2009). "PiSQRD: a web server for decomposing proteins into quasi-rigid dynamical domains." (<http://bioinformatics.oxfordjournals.org/cgi/content/abstract/btp512>). *Bioinformatics* **25** (20): 2743–4. doi:10.1093/bioinformatics/btp512. PMID 19696046. .
- [79] Barclay A (2003). "Membrane proteins with immunoglobulin-like domains--a master superfamily of interaction molecules". *Semin Immunol* **15** (4): 215–23. doi:10.1016/S1044-5323(03)00047-2. PMID 14690046.
- [80] <http://www.compbio.dundee.ac.uk/3Dee/>
- [81] <http://www.cathdb.info/>
- [82] http://ekhidna.biocenter.helsinki.fi/dali_server/
- [83] <http://scop.mrc-lmb.cam.ac.uk/scop>
- [84] http://pawsonlab.mshri.on.ca/index.php?option=com_content&task=view&id=30&Itemid=63/
- [85] http://nashlab.bsd.uchicago.edu/index.php?option=com_content&task=section&id=4&Itemid=29
- [86] <http://realm.sdsc.edu/pdomains/>
- [87] <http://www.ebi.ac.uk/interpro>
- [88] <http://pfam.sanger.ac.uk/>
- [89] <http://www.expasy.org/prosite/>
- [90] <http://prodom.prabi.fr>
- [91] <http://SMART.embl-heidelberg.de>
- [92] <http://www.ncbi.nlm.nih.gov/80/Structure/cdd/cdd.shtml>
- [93] <http://supfam.org/SUPERFAMILY>
- [94] <http://www.sanger.ac.uk/Software/Pfam/browse/clans.shtml#159>
- [95] <http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=10592235>
- [96] <http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=11125048>
- [97] <http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=10592241>
- [98] <http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=10592234>
- [99] <http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=9847202>
- [100] <http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=11125040>

Nucleic Acids

A **nucleic acid** is a macromolecule composed of chains of monomeric nucleotides. In biochemistry these molecules carry genetic information or form structures within cells. The most common nucleic acids are **deoxyribonucleic acid** (DNA) and **ribonucleic acid** (RNA). Nucleic acids are universal in living things, as they are found in all cells and viruses. Nucleic acids were first discovered by Friedrich Miescher in 1871.

Artificial nucleic acids include **peptide nucleic acid** (PNA), **Morpholino** and **locked nucleic acid** (LNA), as well as **glycol nucleic acid** (GNA) and **threose nucleic acid** (TNA). Each of these is distinguished from naturally-occurring DNA or RNA by changes to the backbone of the molecule.

Chemical structure

The term "nucleic acid" is the generic name for a family of biopolymers, named for their role in the cell nucleus. It was later discovered that some nucleic acids are exclusive of the mitochondrion (e.g. Mitochondrial DNA). The monomers from which nucleic acids are constructed are called nucleotides. Nucleic acids are linear, unbranched polymers of nucleotides.

Each nucleotide consists of three components: a nitrogenous heterocyclic base, which is either a purine or a pyrimidine; a pentose sugar; and a phosphate group. Nucleic acid types differ in the structure of the sugar in their nucleotides - DNA contains 2-deoxyribose while RNA contains ribose (where the only difference is the presence of a hydroxyl group). Also, the nitrogenous bases found in the two nucleic acid types are different: adenine, cytosine, and guanine are found in both RNA and DNA, while thymine only occurs in DNA and uracil only occurs in RNA. Other rare nucleic acid bases can occur, for example inosine in strands of mature transfer RNA.

Nucleic acids are usually either single-stranded or double-stranded, though structures with three or more strands can form. A double-stranded nucleic acid consists of two single-stranded nucleic acids held together by hydrogen bonds, such as in the DNA double helix. In contrast, RNA is usually single-stranded, but any given strand may fold back upon itself to form secondary structure as in tRNA and rRNA. Within cells, DNA is usually double-stranded, though some viruses have single-stranded DNA as their genome. Retroviruses have single-stranded RNA as their genome.

The sugars and phosphates in nucleic acids are connected to each other in an alternating chain, linked by shared oxygens, forming a phosphodiester bond. In conventional nomenclature, the carbons to which the phosphate groups attach are the 3' end and the 5' end carbons of the sugar. This gives nucleic acids polarity. The bases extend from a glycosidic linkage to the 1' carbon of the pentose sugar ring. Bases are joined through N-1 of pyrimidines and N-9 of purines to 1' carbon of ribose through N-β glycosyl bond.

Types of nucleic acids

Ribonucleic acid

Ribonucleic acid, or RNA, is a nucleic acid polymer consisting of nucleotide monomers, which plays several important roles in the processes of transcribing genetic information from deoxyribonucleic acid (DNA) into proteins. RNA acts as a messenger between DNA and the protein synthesis complexes known as ribosomes, forms vital portions of ribosomes, and serves as an essential carrier molecule for amino acids to be used in protein synthesis. The three types of RNA include tRNA (transfer), mRNA (messenger) and rRNA (ribosomal).

Deoxyribonucleic acid

Deoxyribonucleic acid is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms. The main role of DNA molecules is the long-term storage of information and DNA is often compared to a set of blueprints, since it contains the instructions needed to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

DNA is made of four types of nucleotides, containing different nucleobases: the pyrimidines cytosine and thymine, and the purines guanine and adenine. The nucleotides are attached to each other in a chain by bonds between their sugar and phosphate groups, forming a sugar-phosphate backbone. Two of these chains are held together by hydrogen bonding between complementary bases; the chains coil around each other, forming the DNA double helix.

Nucleic acid components

Nucleobases

Nucleobases are heterocyclic aromatic organic compounds containing nitrogen atoms. Nucleobases are the parts of RNA and DNA involved in base pairing. Cytosine, guanine, adenine, thymine are found predominantly in DNA, while in RNA uracil replaces thymine. These are abbreviated as C, G, A, T, U, respectively.

Nucleobases are complementary, and when forming base pairs, must always join accordingly: cytosine-guanine, adenine-thymine (adenine-uracil when RNA). The strength of the interaction between cytosine and guanine is stronger than between adenine and thymine because the former pair has three hydrogen bonds joining them while the latter pair have only two. Thus, the higher the GC content of double-stranded DNA, the more stable the molecule and the higher the melting temperature.

Two main nucleobase classes exist, named for the molecule which forms their skeleton. These are the double-ringed purines and single-ringed pyrimidines. Adenine and guanine are purines (abbreviated as R), while cytosine, thymine, and uracil are all pyrimidines (abbreviated as Y).

Hypoxanthine and xanthine are mutant forms of adenine and guanine, respectively, created through mutagen presence, through deamination (replacement of the amine-group with a hydroxyl-group). These are abbreviated HX and X

Nucleosides

Nucleosides are glycosylamines made by attaching a nucleobase (often referred to simply as bases) to a ribose or deoxyribose (sugar) ring. In short, a nucleoside is a base linked to sugar. The names derive from the nucleobase names. The nucleosides commonly occurring in DNA and RNA include cytidine, uridine, adenosine, guanosine and thymidine. When a phosphate is added to a nucleoside (by phosphorylated by a specific kinase enzyme), a nucleotide is produced. Nucleoside analogues, such as acyclovir, are sometimes used as antiviral agents.

Nucleotides and deoxynucleotides

A nucleotide consists of a nucleoside and one phosphate group. Nucleotides are the monomers of RNA and DNA, as well as forming the structural units of several important cofactors - CoA, flavin adenine dinucleotide, flavin mononucleotide, adenosine triphosphate and nicotinamide adenine dinucleotide phosphate. In the cell nucleotides play important roles in metabolism, and signaling.

Nucleotides are named after the nucleoside on which they are based, in conjunction with the number of phosphates they contain, for example:

- Adenine bonded to ribose forms the nucleoside adenosine.
- Adenosine bonded to a phosphate forms adenosine monophosphate.
- As phosphates are added, adenosine diphosphate and adenosine triphosphate are formed, in sequence.

See also

- Nucleic acid methods
- Nucleic acid simulations

References

- Wolfram Saenger, *Principles of Nucleic Acid Structure*, 1984, Springer-Verlag New York Inc.
- Keith Roberts, Martin Raff, Bruce Alberts, Peter Walter, Julian Lewis and Alexander Johnson, *Molecular Biology of the Cell* 4th Edition, Routledge, March, 2002, hardcover, 1616 pages, 7.6 pounds, ISBN 0-8153-3218-1

External links

- Interview with Aaron Klug, Nobel Laureate for structural elucidation of biologically important nucleic-acid protein complexes ^[1] provided by the Vega Science Trust.
- Nucleic Acid Research Journal ^[2]

References

[1] <http://www.vega.org.uk/video/programme/122>

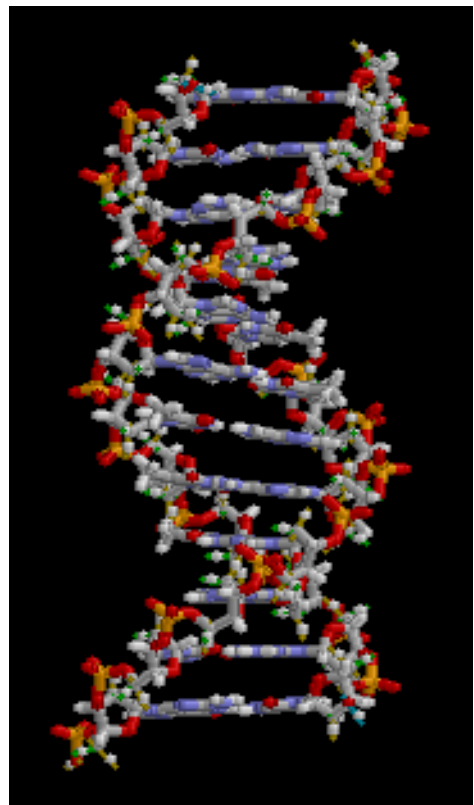
[2] <http://nar.oxfordjournals.org/>

DNA

Deoxyribonucleic acid ( /diˈɒksɪˈraɪboʊnuːkliːk ˈæsɪd/ Wikipedia:Media helpFile:en-us-Deoxyribonucleic_acid.ogg)

(**DNA**) is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms and some viruses. The main role of DNA molecules is the long-term storage of information. DNA is often compared to a set of blueprints or a recipe, or a code, since it contains the instructions needed to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

Chemically, DNA consists of two long polymers of simple units called nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds. These two strands run in opposite directions to each other and are therefore anti-parallel. Attached to each sugar is one of four types of molecules called bases. It is the sequence of these four bases along the backbone that encodes information. This information is read using the genetic code, which specifies the sequence of the amino acids within proteins. The code is read by copying stretches of DNA into the related nucleic acid RNA, in a process called transcription.



The structure of part of a DNA double helix

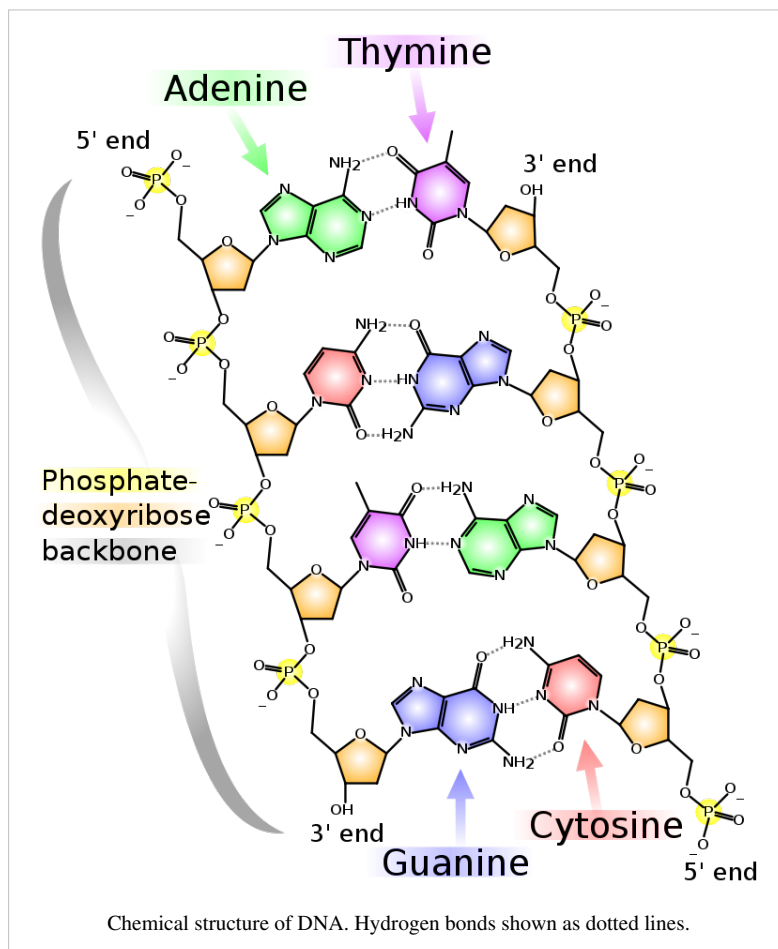
Within cells, DNA is organized into long structures called chromosomes. These chromosomes are duplicated before cells divide, in a process called DNA replication. Eukaryotic organisms (animals, plants, fungi, and protists) store most of their DNA inside the cell nucleus and some of their DNA in organelles, such as mitochondria or chloroplasts.^[1] In contrast, prokaryotes (bacteria and archaea) store their DNA only in the cytoplasm. Within the chromosomes, chromatin proteins such as histones compact and organize DNA. These compact structures guide the interactions between DNA and other proteins, helping control which parts of the DNA are transcribed.

Properties

DNA is a long polymer made from repeating units called nucleotides.^{[2] [3] [4]} The DNA chain is 22 to 26 Ångströms wide (2.2 to 2.6 nanometres), and one nucleotide unit is 3.3 Å (0.33 nm) long.^[5] Although each individual repeating unit is very small, DNA polymers can be very large molecules containing millions of nucleotides. For instance, the largest human chromosome, chromosome number 1, is approximately 220 million base pairs long.^[6]

In living organisms, DNA does not usually exist as a single molecule, but instead as a pair of molecules that are held tightly together.^{[7] [8]} These two long strands entwine like vines, in the shape of a double helix. The nucleotide repeats contain both the segment of the backbone of the molecule, which holds the chain together, and a base, which interacts with the other DNA strand in the helix. A base linked to a sugar is called a nucleoside and a base linked to a sugar and one or more phosphate groups is called a nucleotide. If multiple nucleotides are linked together, as in DNA, this polymer is called a polynucleotide.^[9]

The backbone of the DNA strand is made from alternating phosphate and sugar residues.^[10] The sugar in DNA is 2-deoxyribose, which is a pentose (five-carbon) sugar. The sugars are joined together by phosphate groups that form phosphodiester bonds between the third and fifth carbon atoms of adjacent sugar rings. These asymmetric bonds mean a strand of DNA has a direction. In a double helix the direction of the nucleotides in one strand is opposite to their direction in the other strand: the strands are *antiparallel*. The asymmetric ends of DNA strands are called the 5' (*five prime*) and 3' (*three prime*) ends, with the 5' end having a terminal phosphate group and the 3' end a terminal hydroxyl group. One major difference between DNA and RNA is the sugar, with the 2-deoxyribose in DNA being replaced by the alternative pentose sugar ribose in RNA.^[8]



The DNA double helix is stabilized by hydrogen bonds between the bases attached to the two strands. The four bases found in DNA are adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T). These four bases are attached to the sugar/phosphate to form the complete nucleotide, as shown for adenosine monophosphate.

These bases are classified into two types; adenine and guanine are fused five- and six-membered heterocyclic compounds called purines, while cytosine and thymine are six-membered rings called pyrimidines.^[8] A fifth pyrimidine base, called uracil (U), usually takes the place of thymine in RNA and differs from thymine by lacking a methyl group on its ring. Uracil is not usually found in DNA, occurring only as a breakdown product of cytosine. In addition to RNA and DNA, a large number of artificial nucleic acid analogues have also been created to study the properties of nucleic acids, or for use in biotechnology.^[12]

Grooves

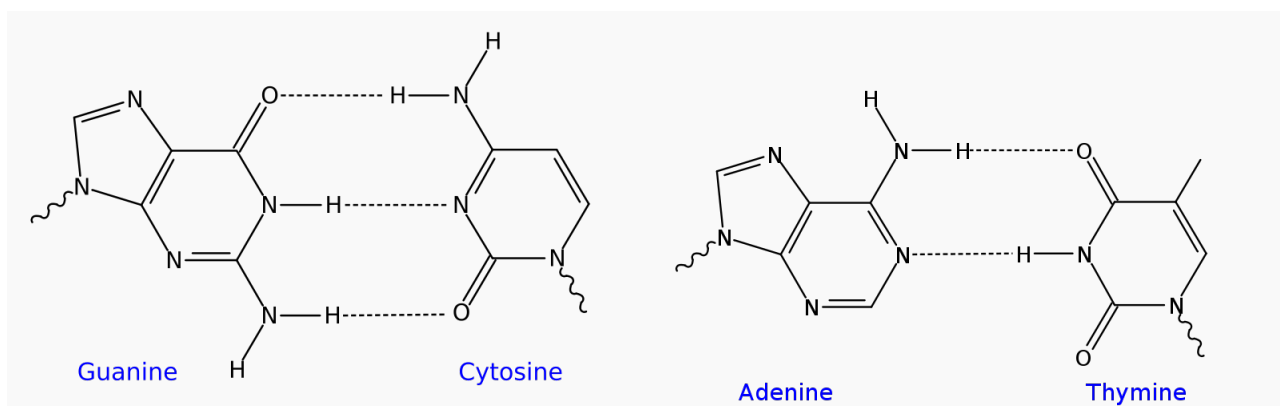
Twin helical strands form the DNA backbone. Another double helix may be found by tracing the spaces, or grooves, between the strands. These voids are adjacent to the base pairs and may provide a binding site. As the strands are not directly opposite each other, the grooves are unequally sized. One groove, the major groove, is 22 Å wide and the other, the minor groove, is 12 Å wide.^[13] The narrowness of the minor groove means that the edges of the bases are more accessible in the major groove. As a result, proteins like transcription factors that can bind to specific sequences in double-stranded DNA usually make contacts to the sides of the bases exposed in the major groove.^[14] This situation varies in unusual conformations of DNA within the cell (*see below*), but the major and minor grooves are always named to reflect the differences in size that would be seen if the DNA is twisted back into the ordinary B form.

Base pairing

Each type of base on one strand forms a bond with just one type of base on the other strand. This is called complementary base pairing. Here, purines form hydrogen bonds to pyrimidines, with A bonding only to T, and C bonding only to G. This arrangement of two nucleotides binding together across the double helix is called a base pair. As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily. The two strands of DNA in a double helix can therefore be pulled apart like a zipper, either by a mechanical force or high temperature.^[15] As a result of this complementarity, all the information in the double-stranded sequence of a DNA helix is duplicated on each strand, which is vital in DNA replication. Indeed, this reversible and specific interaction between complementary base pairs is critical for all the functions of DNA in living organisms.^[3]



A section of DNA. The bases lie horizontally between the two spiraling strands.^[11] Animated version at File:DNA orbit animated.gif.



Top, a **GC** base pair with three hydrogen bonds. Bottom, an **AT** base pair with two hydrogen bonds. Non-covalent hydrogen bonds between the pairs are shown as dashed lines.

The two types of base pairs form different numbers of hydrogen bonds, AT forming two hydrogen bonds, and GC forming three hydrogen bonds (see figures, left). DNA with high GC-content is more stable than DNA with low GC-content, but contrary to popular belief, this is not due to the extra hydrogen bond of a GC base pair but rather the contribution of stacking interactions (hydrogen bonding merely provides specificity of the pairing, not stability).^[16] As a result, it is both the percentage of GC base pairs and the overall length of a DNA double helix that determine the strength of the association between the two strands of DNA. Long DNA helices with a high GC content have stronger-interacting strands, while short helices with high AT content have weaker-interacting strands.^[17] In biology, parts of the DNA double helix that need to separate easily, such as the TATAAT Pribnow box in some promoters, tend to have a high AT content, making the strands easier to pull apart.^[18] In the laboratory, the strength of this interaction can be measured by finding the temperature required to break the hydrogen bonds, their melting temperature (also called T_m value). When all the base pairs in a DNA double helix melt, the strands separate and exist in solution as two entirely independent molecules. These single-stranded DNA molecules have no single common shape, but some conformations are more stable than others.^[19]

Sense and antisense

A DNA sequence is called "sense" if its sequence is the same as that of a messenger RNA copy that is translated into protein.^[20] The sequence on the opposite strand is called the "antisense" sequence. Both sense and antisense sequences can exist on different parts of the same strand of DNA (i.e. both strands contain both sense and antisense sequences). In both prokaryotes and eukaryotes, antisense RNA sequences are produced, but the functions of these RNAs are not entirely clear.^[21] One proposal is that antisense RNAs are involved in regulating gene expression through RNA-RNA base pairing.^[22]

A few DNA sequences in prokaryotes and eukaryotes, and more in plasmids and viruses, blur the distinction between sense and antisense strands by having overlapping genes.^[23] In these cases, some DNA sequences do double duty, encoding one protein when read along one strand, and a second protein when read in the opposite direction along the other strand. In bacteria, this overlap may be involved in the regulation of gene transcription,^[24] while in viruses, overlapping genes increase the amount of information that can be encoded within the small viral genome.^[25]

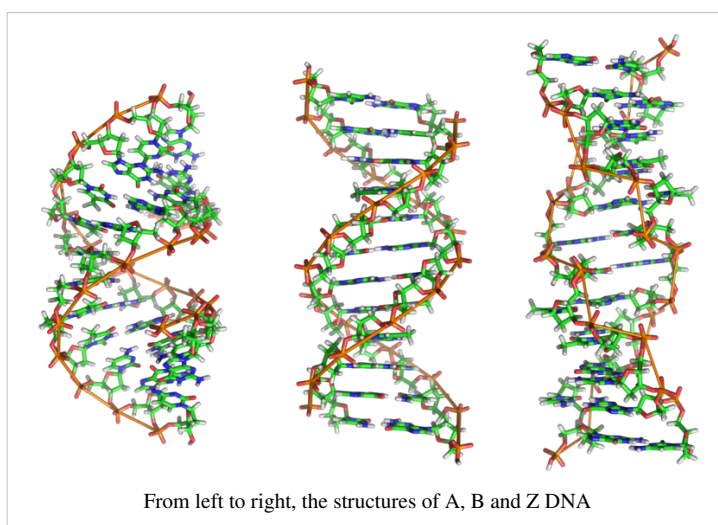
Supercoiling

DNA can be twisted like a rope in a process called DNA supercoiling. With DNA in its "relaxed" state, a strand usually circles the axis of the double helix once every 10.4 base pairs, but if the DNA is twisted the strands become more tightly or more loosely wound.^[26] If the DNA is twisted in the direction of the helix, this is positive supercoiling, and the bases are held more tightly together. If they are twisted in the opposite direction, this is negative supercoiling, and the bases come apart more easily. In nature, most DNA has slight negative supercoiling that is introduced by enzymes called topoisomerases.^[27] These enzymes are also needed to relieve the twisting

stresses introduced into DNA strands during processes such as transcription and DNA replication.^[28]

Alternate DNA structures

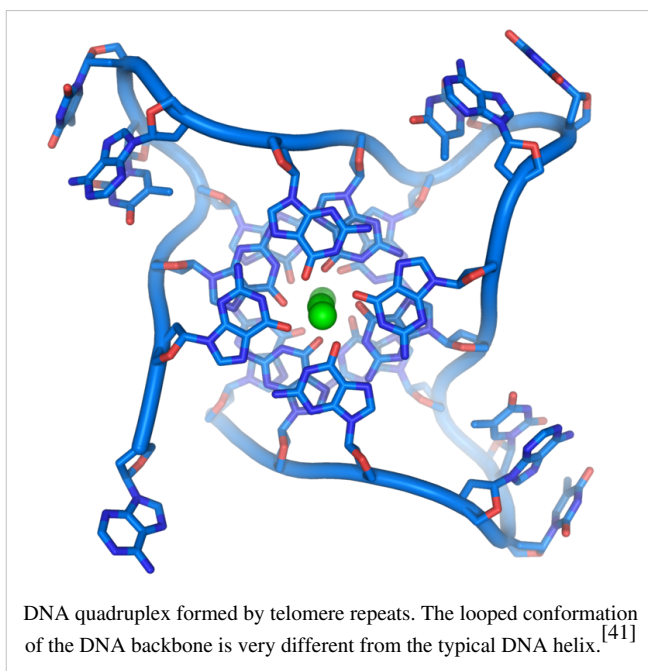
DNA exists in many possible conformations that include A-DNA, B-DNA, and Z-DNA forms, although, only B-DNA and Z-DNA have been directly observed in functional organisms.^[10] The conformation that DNA adopts depends on the hydration level, DNA sequence, the amount and direction of supercoiling, chemical modifications of the bases, the type and concentration of metal ions, as well as the presence of polyamines in solution.^[29]



The first published reports of A-DNA X-ray diffraction patterns— and also B-DNA used analyses based on Patterson transforms that provided only a limited amount of structural information for oriented fibers of DNA.^{[30] [31]} An alternate analysis was then proposed by Wilkins *et al.*, in 1953, for the *in vivo* B-DNA X-ray diffraction/scattering patterns of highly hydrated DNA fibers in terms of squares of Bessel functions.^[32] In the same journal, Watson and Crick presented their molecular modeling analysis of the DNA X-ray diffraction patterns to suggest that the structure was a double-helix.^[7]

Although the 'B-DNA form' is most common under the conditions found in cells,^[33] it is not a well-defined conformation but a family of related DNA conformations^[34] that occur at the high hydration levels present in living cells. Their corresponding X-ray diffraction and scattering patterns are characteristic of molecular paracrystals with a significant degree of disorder.^{[35] [36]}

Compared to B-DNA, the A-DNA form is a wider right-handed spiral, with a shallow, wide minor groove and a narrower, deeper major groove. The A form occurs under non-physiological conditions in partially dehydrated samples of DNA, while in the cell it may be produced in hybrid pairings of DNA and RNA strands, as well as in enzyme-DNA complexes.^{[37] [38]} Segments of DNA where the bases have been chemically modified by methylation may undergo a larger change in conformation and adopt the Z form. Here, the strands turn about the helical axis in a left-handed spiral, the opposite of the more common B form.^[39] These unusual structures can be recognized by specific Z-DNA binding proteins and may be involved in the regulation of transcription.^[40]



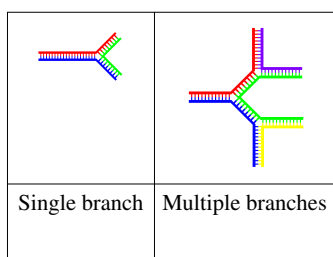
Quadruplex structures

At the ends of the linear chromosomes are specialized regions of DNA called telomeres. The main function of these regions is to allow the cell to replicate chromosome ends using the enzyme telomerase, as the enzymes that normally replicate DNA cannot copy the extreme 3' ends of chromosomes.^[42] These specialized chromosome caps also help protect the DNA ends, and stop the DNA repair systems in the cell from treating them as damage to be corrected.^[43] In human cells, telomeres are usually lengths of single-stranded DNA containing several thousand repeats of a simple TTAGGG sequence.^[44]

These guanine-rich sequences may stabilize chromosome ends by forming structures of stacked sets of four-base units, rather than the usual base

pairs found in other DNA molecules. Here, four guanine bases form a flat plate and these flat four-base units then stack on top of each other, to form a stable *G-quadruplex* structure.^[45] These structures are stabilized by hydrogen bonding between the edges of the bases and chelation of a metal ion in the centre of each four-base unit.^[46] Other structures can also be formed, with the central set of four bases coming from either a single strand folded around the bases, or several different parallel strands, each contributing one base to the central structure.

In addition to these stacked structures, telomeres also form large loop structures called telomere loops, or T-loops. Here, the single-stranded DNA curls around in a long circle stabilized by telomere-binding proteins.^[47] At the very end of the T-loop, the single-stranded telomere DNA is held onto a region of double-stranded DNA by the telomere strand disrupting the double-helical DNA and base pairing to one of the two strands. This triple-stranded structure is called a displacement loop or D-loop.^[45]

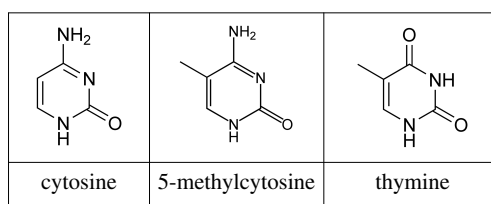


Branched DNA can form networks containing multiple branches.

Branched DNA

In DNA fraying occurs when non-complementary regions exist at the end of an otherwise complementary double-strand of DNA. However, branched DNA can occur if a third strand of DNA is introduced and contains adjoining regions able to hybridize with the frayed regions of the pre-existing double-strand. Although the simplest example of branched DNA involves only three strands of DNA, complexes involving additional strands and multiple branches are also possible.^[48] Branched DNA can be used in nanotechnology to construct geometric shapes, see the section on uses in technology below.

Chemical modifications



Structure of cytosine with and without the 5-methyl group. Deamination converts 5-methylcytosine into thymine.

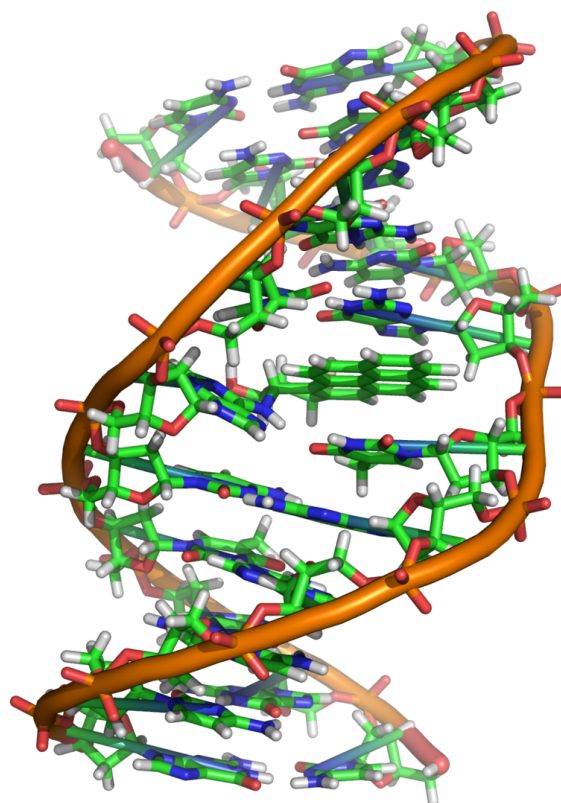
Base modifications

The expression of genes is influenced by how the DNA is packaged in chromosomes, in a structure called chromatin. Base modifications can be involved in packaging, with regions that have low or no gene expression usually containing high levels of methylation of cytosine bases. For example, cytosine methylation, produces 5-methylcytosine, which is important for X-chromosome inactivation.^[49] The average level of methylation varies between organisms - the worm *Caenorhabditis elegans* lacks cytosine methylation, while vertebrates have higher levels, with up to 1% of their DNA containing 5-methylcytosine.^[50] Despite the importance of 5-methylcytosine, it can deaminate to leave a thymine base, methylated cytosines are therefore particularly prone to mutations.^[51] Other base modifications include adenine methylation in bacteria, the presence of 5-hydroxymethylcytosine in the brain,^[52] and the glycosylation of uracil to produce the "J-base" in kinetoplasts.^[53] ^[54]

Damage

DNA can be damaged by many sorts of mutagens, which change the DNA sequence. Mutagens include oxidizing agents, alkylating agents and also high-energy electromagnetic radiation such as ultraviolet light and X-rays. The type of DNA damage produced depends on the type of mutagen. For example, UV light can damage DNA by producing thymine dimers, which are cross-links between pyrimidine bases.^[56] On the other hand, oxidants such as free radicals or hydrogen peroxide produce multiple forms of damage, including base modifications, particularly of guanosine, and double-strand breaks.^[57] A typical human cell contains about 150,000 bases that have suffered oxidative damage.^[58] Of these oxidative lesions, the most dangerous are double-strand breaks, as these are difficult to repair and can produce point mutations, insertions and deletions from the DNA sequence, as well as chromosomal translocations.^[59]

Many mutagens fit into the space between two adjacent base pairs, this is called *intercalating*. Most intercalators are aromatic and planar molecules, and include Ethidium bromide, daunomycin, and



A covalent adduct between benzo[*a*]pyrene, the major mutagen in tobacco smoke, and DNA^[55]

doxorubicin. In order for an intercalator to fit between base pairs, the bases must separate, distorting the DNA strands by unwinding of the double helix. This inhibits both transcription and DNA replication, causing toxicity and mutations. As a result, DNA intercalators are often carcinogens, and Benzo[*a*]pyrene diol epoxide, acridines, aflatoxin and ethidium bromide are well-known examples.^{[60] [61] [62]} Nevertheless, due to their ability to inhibit DNA transcription and replication, other similar toxins are also used in chemotherapy to inhibit rapidly growing cancer cells.^[63]

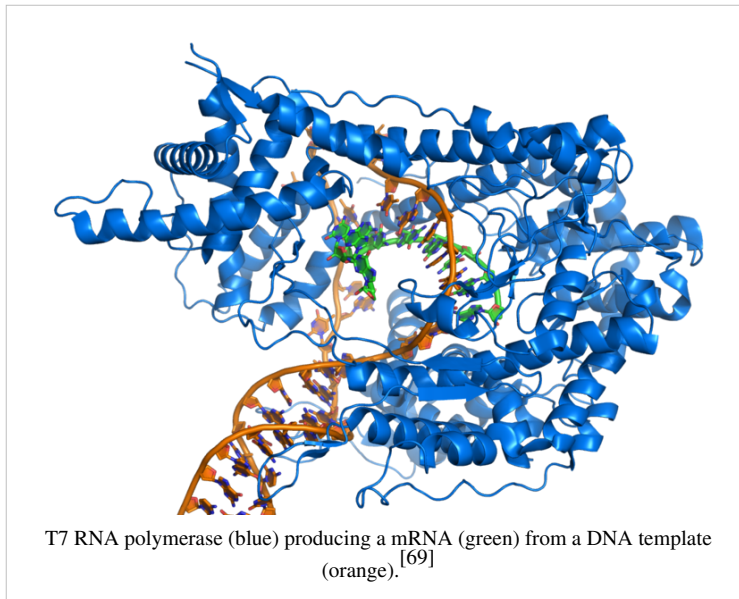
Biological functions

DNA usually occurs as linear chromosomes in eukaryotes, and circular chromosomes in prokaryotes. The set of chromosomes in a cell makes up its genome; the human genome has approximately 3 billion base pairs of DNA arranged into 46 chromosomes.^[64] The information carried by DNA is held in the sequence of pieces of DNA called genes. Transmission of genetic information in genes is achieved via complementary base pairing. For example, in transcription, when a cell uses the information in a gene, the DNA sequence is copied into a complementary RNA sequence through the attraction between the DNA and the correct RNA nucleotides. Usually, this RNA copy is then used to make a matching protein sequence in a process called translation which depends on the same interaction between RNA nucleotides. Alternatively, a cell may simply copy its genetic information in a process called DNA replication. The details of these functions are covered in other articles; here we focus on the interactions between DNA and other molecules that mediate the function of the genome.

Genes and genomes

Genomic DNA is located in the cell nucleus of eukaryotes, as well as small amounts in mitochondria and chloroplasts. In prokaryotes, the DNA is held within an irregularly shaped body in the cytoplasm called the nucleoid.^[65] The genetic information in a genome is held within genes, and the complete set of this information in an organism is called its genotype. A gene is a unit of heredity and is a region of DNA that influences a particular characteristic in an organism. Genes contain an open reading frame that can be transcribed, as well as regulatory sequences such as promoters and enhancers, which control the transcription of the open reading frame.

In many species, only a small fraction of the total sequence of the genome encodes protein. For example, only about 1.5% of the human genome consists of protein-coding exons, with over 50% of human DNA consisting of non-coding repetitive sequences.^[66] The reasons for the presence of so much non-coding DNA in eukaryotic genomes and the extraordinary differences in genome size, or *C-value*, among species represent a long-standing puzzle known as the "C-value enigma."^[67] However, DNA sequences that do not code protein may still encode functional non-coding RNA molecules, which are involved in the regulation of gene expression.^[68]



Some non-coding DNA sequences play structural roles in chromosomes. Telomeres and centromeres typically contain few genes, but are important for the function and stability of chromosomes.^{[43] [70]} An abundant form of non-coding DNA in humans are pseudogenes, which are copies of genes that have been disabled by mutation.^[71] These sequences are usually just molecular fossils, although they can occasionally serve as raw genetic material for the creation of new genes through the process of gene duplication and divergence.^[72]

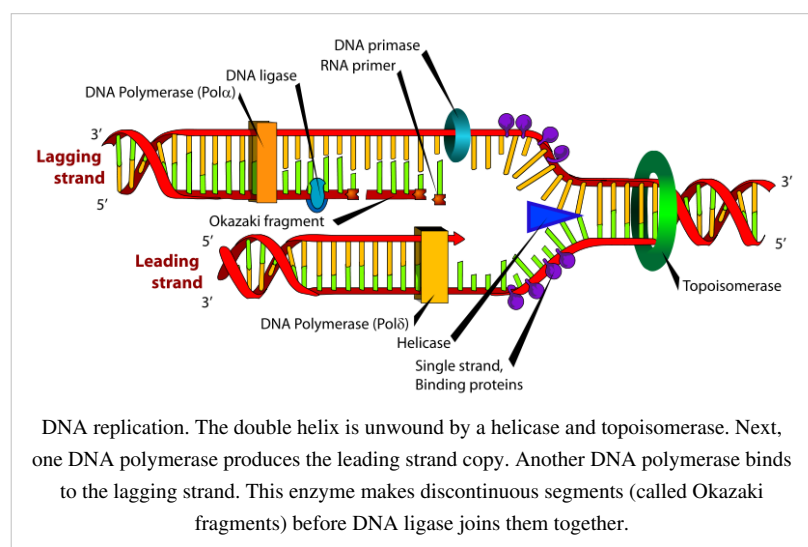
Transcription and translation

A gene is a sequence of DNA that contains genetic information and can influence the phenotype of an organism. Within a gene, the sequence of bases along a DNA strand defines a messenger RNA sequence, which then defines one or more protein sequences. The relationship between the nucleotide sequences of genes and the amino-acid sequences of proteins is determined by the rules of translation, known collectively as the genetic code. The genetic code consists of three-letter 'words' called *codons* formed from a sequence of three nucleotides (e.g. ACT, CAG, TTT).

In transcription, the codons of a gene are copied into messenger RNA by RNA polymerase. This RNA copy is then decoded by a ribosome that reads the RNA sequence by base-pairing the messenger RNA to transfer RNA, which carries amino acids. Since there are 4 bases in 3-letter combinations, there are 64 possible codons (4^3 combinations). These encode the twenty standard amino acids, giving most amino acids more than one possible codon. There are also three 'stop' or 'nonsense' codons signifying the end of the coding region; these are the TAA, TGA and TAG codons.

Replication

Cell division is essential for an organism to grow, but when a cell divides it must replicate the DNA in its genome so that the two daughter cells have the same genetic information as their parent. The double-stranded structure of DNA provides a simple mechanism for DNA replication. Here, the two strands are separated and then each strand's complementary DNA sequence is recreated by an enzyme called DNA polymerase. This enzyme makes the complementary strand by



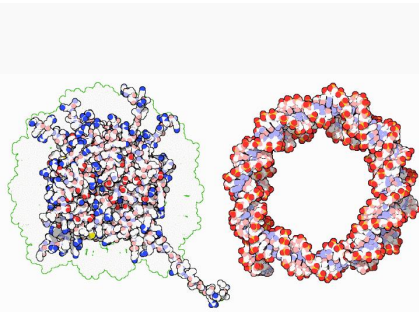
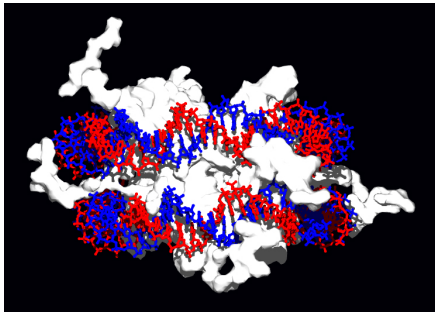
finding the correct base through complementary base pairing, and bonding it onto the original strand. As DNA polymerases can only extend a DNA strand in a 5' to 3' direction, different mechanisms are used to copy the

antiparallel strands of the double helix.^[73] In this way, the base on the old strand dictates which base appears on the new strand, and the cell ends up with a perfect copy of its DNA.

Interactions with proteins

All the functions of DNA depend on interactions with proteins. These protein interactions can be non-specific, or the protein can bind specifically to a single DNA sequence. Enzymes can also bind to DNA and of these, the polymerases that copy the DNA base sequence in transcription and DNA replication are particularly important.

DNA-binding proteins



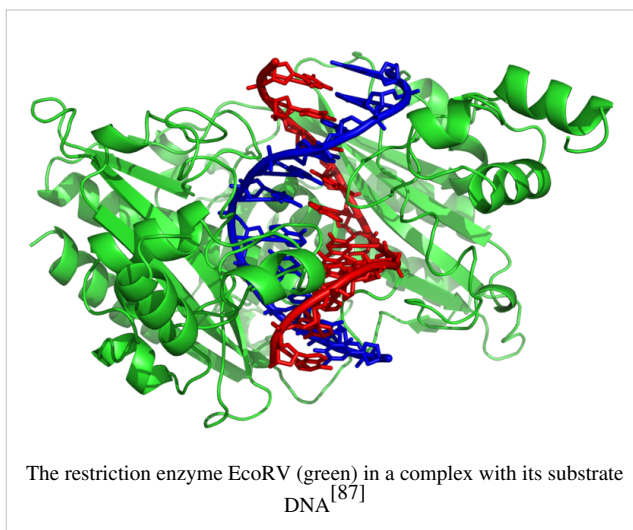
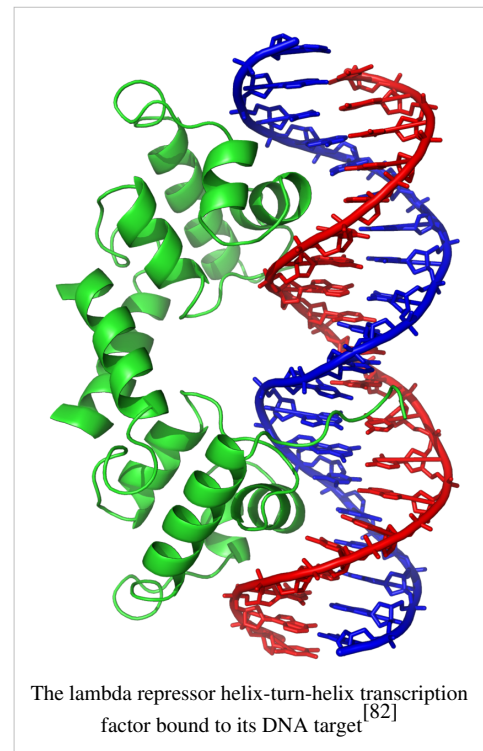
Interaction of DNA with histones (shown in white, top). These proteins' basic amino acids (below left, blue) bind to the acidic phosphate groups on DNA (below right, red).

Structural proteins that bind DNA are well-understood examples of non-specific DNA-protein interactions. Within chromosomes, DNA is held in complexes with structural proteins. These proteins organize the DNA into a compact structure called chromatin. In eukaryotes this structure involves DNA binding to a complex of small basic proteins called histones, while in prokaryotes multiple types of proteins are involved.^{[74] [75]} The histones form a disk-shaped complex called a nucleosome, which contains two complete turns of double-stranded DNA wrapped around its surface. These non-specific interactions are formed through basic residues in the histones making ionic bonds to the acidic sugar-phosphate backbone of the DNA, and are therefore largely independent of the base sequence.^[76] Chemical modifications of these basic amino acid residues include methylation, phosphorylation and acetylation.^[77] These chemical changes alter the strength of the interaction between the DNA and the histones, making the DNA more or less accessible to transcription factors and changing the rate of transcription.^[78] Other non-specific DNA-binding proteins in chromatin include the high-mobility group proteins, which bind to bent or distorted DNA.^[79] These proteins are important in bending arrays of nucleosomes and arranging them into the larger structures that make up chromosomes.^[80]

A distinct group of DNA-binding proteins are the DNA-binding proteins that specifically bind single-stranded DNA. In humans, replication protein A is the best-understood member of this family and is used in processes where the double helix is separated, including DNA replication, recombination and DNA repair.^[81] These binding proteins seem to stabilize single-stranded DNA and protect it from forming stem-loops or being degraded by nucleases.

In contrast, other proteins have evolved to bind to particular DNA sequences. The most intensively studied of these are the various transcription factors, which are proteins that regulate transcription. Each transcription factor binds to one particular set of DNA sequences and activates or inhibits the transcription of genes that have these sequences close to their promoters. The transcription factors do this in two ways. Firstly, they can bind the RNA polymerase responsible for transcription, either directly or through other mediator proteins; this locates the polymerase at the promoter and allows it to begin transcription.^[83] Alternatively, transcription factors can bind enzymes that modify the histones at the promoter; this will change the accessibility of the DNA template to the polymerase.^[84]

As these DNA targets can occur throughout an organism's genome, changes in the activity of one type of transcription factor can affect thousands of genes.^[85] Consequently, these proteins are often the targets of the signal transduction processes that control responses to environmental changes or cellular differentiation and development. The specificity of these transcription factors' interactions with DNA come from the proteins making multiple contacts to the edges of the DNA bases, allowing them to "read" the DNA sequence. Most of these base-interactions are made in the major groove, where the bases are most accessible.^[86]



DNA-modifying enzymes

Nucleases and ligases

Nucleases are enzymes that cut DNA strands by catalyzing the hydrolysis of the phosphodiester bonds. Nucleases that hydrolyse nucleotides from the ends of DNA strands are called exonucleases, while endonucleases cut within strands. The most frequently used nucleases in molecular biology are the restriction endonucleases, which cut DNA at specific sequences. For instance, the EcoRV enzyme shown to the left recognizes the 6-base sequence 5'-GAT|ATC-3' and makes a cut at the vertical line. In nature, these enzymes protect bacteria against phage infection by

digesting the phage DNA when it enters the bacterial cell, acting as part of the restriction modification system.^[88] In technology, these sequence-specific nucleases are used in molecular cloning and DNA fingerprinting.

Enzymes called DNA ligases can rejoin cut or broken DNA strands.^[89] Ligases are particularly important in lagging strand DNA replication, as they join together the short segments of DNA produced at the replication fork into a complete copy of the DNA template. They are also used in DNA repair and genetic recombination.^[89]

Topoisomerases and helicases

Topoisomerases are enzymes with both nuclease and ligase activity. These proteins change the amount of supercoiling in DNA. Some of these enzymes work by cutting the DNA helix and allowing one section to rotate, thereby reducing its level of supercoiling; the enzyme then seals the DNA break.^[27] Other types of these enzymes are capable of cutting one DNA helix and then passing a second strand of DNA through this break, before rejoining the helix.^[90] Topoisomerases are required for many processes involving DNA, such as DNA replication and transcription.^[28]

Helicases are proteins that are a type of molecular motor. They use the chemical energy in nucleoside triphosphates, predominantly ATP, to break hydrogen bonds between bases and unwind the DNA double helix into single strands.^[91] These enzymes are essential for most processes where enzymes need to access the DNA bases.

Polymerases

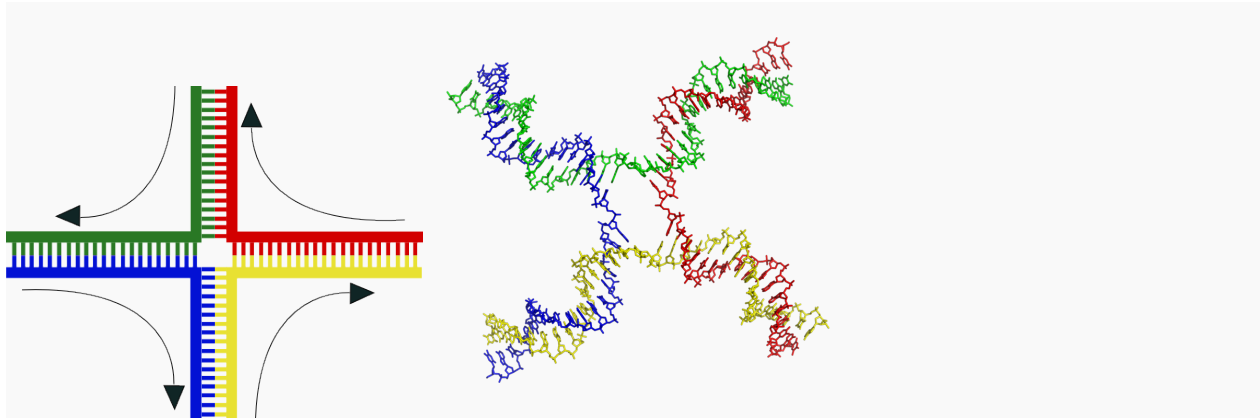
Polymerases are enzymes that synthesize polynucleotide chains from nucleoside triphosphates. The sequence of their products are copies of existing polynucleotide chains - which are called *templates*. These enzymes function by adding nucleotides onto the 3' hydroxyl group of the previous nucleotide in a DNA strand. Consequently, all polymerases work in a 5' to 3' direction.^[92] In the active site of these enzymes, the incoming nucleoside triphosphate base-pairs to the template: this allows polymerases to accurately synthesize the complementary strand of their template. Polymerases are classified according to the type of template that they use.

In DNA replication, a DNA-dependent DNA polymerase makes a copy of a DNA sequence. Accuracy is vital in this process, so many of these polymerases have a proofreading activity. Here, the polymerase recognizes the occasional mistakes in the synthesis reaction by the lack of base pairing between the mismatched nucleotides. If a mismatch is detected, a 3' to 5' exonuclease activity is activated and the incorrect base removed.^[93] In most organisms DNA polymerases function in a large complex called the replisome that contains multiple accessory subunits, such as the DNA clamp or helicases.^[94]

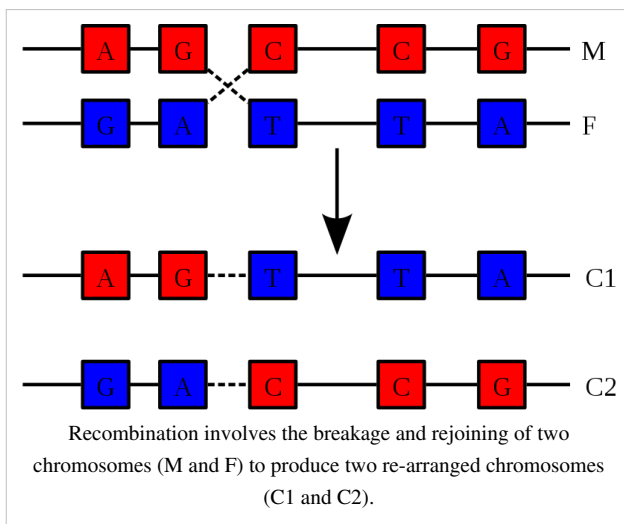
RNA-dependent DNA polymerases are a specialized class of polymerases that copy the sequence of an RNA strand into DNA. They include reverse transcriptase, which is a viral enzyme involved in the infection of cells by retroviruses, and telomerase, which is required for the replication of telomeres.^[42] ^[95] Telomerase is an unusual polymerase because it contains its own RNA template as part of its structure.^[43]

Transcription is carried out by a DNA-dependent RNA polymerase that copies the sequence of a DNA strand into RNA. To begin transcribing a gene, the RNA polymerase binds to a sequence of DNA called a promoter and separates the DNA strands. It then copies the gene sequence into a messenger RNA transcript until it reaches a region of DNA called the terminator, where it halts and detaches from the DNA. As with human DNA-dependent DNA polymerases, RNA polymerase II, the enzyme that transcribes most of the genes in the human genome, operates as part of a large protein complex with multiple regulatory and accessory subunits.^[96]

Genetic recombination



Structure of the Holliday junction intermediate in genetic recombination. The four separate DNA strands are coloured red, blue, green and yellow.^[97]



A DNA helix usually does not interact with other segments of DNA, and in human cells the different chromosomes even occupy separate areas in the nucleus called "chromosome territories".^[98] This physical separation of different chromosomes is important for the ability of DNA to function as a stable repository for information, as one of the few times chromosomes interact is during chromosomal crossover when they recombine. Chromosomal crossover is when two DNA helices break, swap a section and then rejoin.

Recombination allows chromosomes to exchange genetic information and produces new combinations of genes, which increases the efficiency of natural selection and can be important in the rapid evolution of

new proteins.^[99] Genetic recombination can also be involved in DNA repair, particularly in the cell's response to double-strand breaks.^[100]

The most common form of chromosomal crossover is homologous recombination, where the two chromosomes involved share very similar sequences. Non-homologous recombination can be damaging to cells, as it can produce chromosomal translocations and genetic abnormalities. The recombination reaction is catalyzed by enzymes known as recombinases, such as RAD51.^[101] The first step in recombination is a double-stranded break either caused by an endonuclease or damage to the DNA.^[102] A series of steps catalyzed in part by the recombinase then leads to joining of the two helices by at least one Holliday junction, in which a segment of a single strand in each helix is annealed to the complementary strand in the other helix. The Holliday junction is a tetrahedral junction structure that can be moved along the pair of chromosomes, swapping one strand for another. The recombination reaction is then halted by cleavage of the junction and re-ligation of the released DNA.^[103]

Evolution

DNA contains the genetic information that allows all modern living things to function, grow and reproduce. However, it is unclear how long in the 4-billion-year history of life DNA has performed this function, as it has been proposed that the earliest forms of life may have used RNA as their genetic material.^{[92] [104]} RNA may have acted as the central part of early cell metabolism as it can both transmit genetic information and carry out catalysis as part of ribozymes.^[105] This ancient RNA world where nucleic acid would have been used for both catalysis and genetics may have influenced the evolution of the current genetic code based on four nucleotide bases. This would occur since the number of different bases in such an organism is a trade-off between a small number of bases increasing replication accuracy and a large number of bases increasing the catalytic efficiency of ribozymes.^[106]

Unfortunately, there is no direct evidence of ancient genetic systems, as recovery of DNA from most fossils is impossible. This is because DNA will survive in the environment for less than one million years and slowly degrades into short fragments in solution.^[107] Claims for older DNA have been made, most notably a report of the isolation of a viable bacterium from a salt crystal 250 million years old,^[108] but these claims are controversial.^{[109] [110]}

Uses in technology

Genetic engineering

Methods have been developed to purify DNA from organisms, such as phenol-chloroform extraction and manipulate it in the laboratory, such as restriction digests and the polymerase chain reaction. Modern biology and biochemistry make intensive use of these techniques in recombinant DNA technology. Recombinant DNA is a man-made DNA sequence that has been assembled from other DNA sequences. They can be transformed into organisms in the form of plasmids or in the appropriate format, by using a viral vector.^[111] The genetically modified organisms produced can be used to produce products such as recombinant proteins, used in medical research,^[112] or be grown in agriculture.^{[113] [114]}

Forensics

Forensic scientists can use DNA in blood, semen, skin, saliva or hair found at a crime scene to identify a matching DNA of an individual, such as a perpetrator. This process is called genetic fingerprinting, or more accurately, DNA profiling. In DNA profiling, the lengths of variable sections of repetitive DNA, such as short tandem repeats and minisatellites, are compared between people. This method is usually an extremely reliable technique for identifying a matching DNA.^[115] However, identification can be complicated if the scene is contaminated with DNA from several people.^[116] DNA profiling was developed in 1984 by British geneticist Sir Alec Jeffreys,^[117] and first used in forensic science to convict Colin Pitchfork in the 1988 Enderby murders case.^[118]

People convicted of certain types of crimes may be required to provide a sample of DNA for a database. This has helped investigators solve old cases where only a DNA sample was obtained from the scene. DNA profiling can also be used to identify victims of mass casualty incidents.^[119] On the other hand, many convicted people have been released from prison on the basis of DNA techniques, which were not available when a crime had originally been committed.

Bioinformatics

Bioinformatics involves the manipulation, searching, and data mining of DNA sequence data. The development of techniques to store and search DNA sequences have led to widely applied advances in computer science, especially string searching algorithms, machine learning and database theory.^[120] String searching or matching algorithms, which find an occurrence of a sequence of letters inside a larger sequence of letters, were developed to search for specific sequences of nucleotides.^[121] In other applications such as text editors, even simple algorithms for this

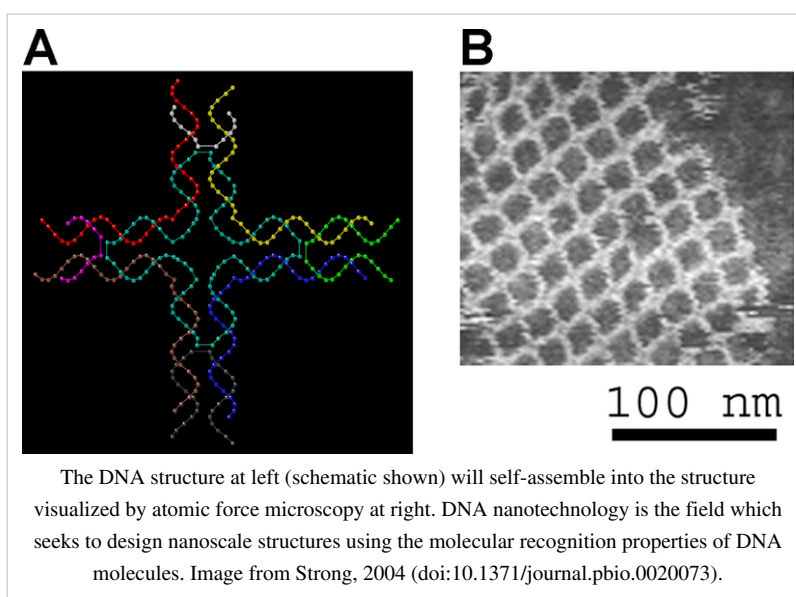
problem usually suffice, but DNA sequences cause these algorithms to exhibit near-worst-case behaviour due to their small number of distinct characters. The related problem of sequence alignment aims to identify homologous sequences and locate the specific mutations that make them distinct. These techniques, especially multiple sequence alignment, are used in studying phylogenetic relationships and protein function.^[122] Data sets representing entire genomes' worth of DNA sequences, such as those produced by the Human Genome Project, are difficult to use without annotations, which label the locations of genes and regulatory elements on each chromosome. Regions of DNA sequence that have the characteristic patterns associated with protein- or RNA-coding genes can be identified by gene finding algorithms, which allow researchers to predict the presence of particular gene products in an organism even before they have been isolated experimentally.^[123]

DNA nanotechnology

DNA nanotechnology uses the unique molecular recognition properties of DNA and other nucleic acids to create self-assembling branched DNA complexes with useful properties.^[124]

DNA is thus used as a structural material rather than as a carrier of biological information. This has led to the creation of two-dimensional periodic lattices (both tile-based as well as using the "DNA origami" method) as well as three-dimensional structures in the shapes of polyhedra.^[125] Nanomechanical devices and algorithmic self-assembly have also been demonstrated,^[126] and

these DNA structures have been used to template the arrangement of other molecules such as gold nanoparticles and streptavidin proteins.^[127]



History and anthropology

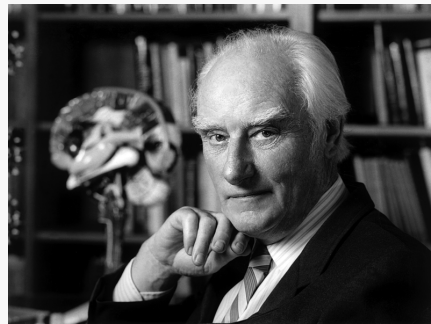
Because DNA collects mutations over time, which are then inherited, it contains historical information and by comparing DNA sequences, geneticists can infer the evolutionary history of organisms, their phylogeny.^[128] This field of phylogenetics is a powerful tool in evolutionary biology. If DNA sequences within a species are compared, population geneticists can learn the history of particular populations. This can be used in studies ranging from ecological genetics to anthropology; for example, DNA evidence is being used to try to identify the Ten Lost Tribes of Israel.^{[129] [130]}

DNA has also been used to look at modern family relationships, such as establishing family relationships between the descendants of Sally Hemings and Thomas Jefferson. This usage is closely related to the use of DNA in criminal investigations detailed above. Indeed, some criminal investigations have been solved when DNA from crime scenes has matched relatives of the guilty individual.^[131]

History of DNA research

DNA was first isolated by the Swiss physician Friedrich Miescher who, in 1869, discovered a microscopic substance in the pus of discarded surgical bandages. As it resided in the nuclei of cells, he called it "nuclein".^[132] In 1919, Phoebus Levene identified the base, sugar and phosphate nucleotide unit.^[133] Levene suggested that DNA consisted of a string of nucleotide units linked together through the phosphate groups. However, Levene thought the chain was short and the bases repeated in a fixed order. In 1937 William Astbury produced the first X-ray diffraction patterns that showed that DNA had a regular structure.^[134]

In 1928, Frederick Griffith discovered that traits of the "smooth" form of the *Pneumococcus* could be transferred to the "rough" form of the same bacteria by mixing killed "smooth" bacteria with the live "rough" form.^[135] This system provided the first clear suggestion that DNA carried genetic information—the Avery-MacLeod-McCarty experiment—when Oswald Avery, along with coworkers Colin MacLeod and Maclyn McCarty, identified DNA as the transforming principle in 1943.^[136] DNA's role in heredity was confirmed in 1952, when Alfred Hershey and Martha Chase in the Hershey-Chase experiment showed that DNA is the genetic material of the T2 phage.^[137]



Francis Crick



Rosalind Franklin



Raymond Gosling

In 1953 James D. Watson and Francis Crick suggested what is now accepted as the first correct double-helix model of DNA structure in the journal *Nature*.^[7] Their double-helix, molecular model of DNA was then based on a single X-ray diffraction image (labeled as "Photo 51")^[138] taken by Rosalind Franklin and Raymond Gosling in May 1952, as well as the information that the DNA bases were paired—also obtained through private communications from Erwin Chargaff in the previous years. Chargaff's rules played a very important role in establishing double-helix configurations for B-DNA as well as A-DNA.

Experimental evidence supporting the Watson and Crick model were published in a series of five articles in the same issue of *Nature*.^[139] Of these, Franklin and Gosling's paper was the first publication of their own X-ray diffraction data and original analysis method that partially supported the Watson and Crick model^{[31] [140]}; this issue also contained an article on DNA structure by Maurice Wilkins and two of his colleagues, whose analysis and *in vivo* B-DNA X-ray patterns also supported the presence *in vivo* of the double-helical DNA configurations as proposed by Crick and Watson for their double-helix molecular model of DNA in the previous two pages of *Nature*.^[32] In 1962, after Franklin's death, Watson, Crick, and Wilkins jointly received the Nobel Prize in Physiology or Medicine.^[141] Unfortunately, Nobel rules of the time allowed only living recipients, but a vigorous debate continues on who should receive credit for the discovery.^[142]

In an influential presentation in 1957, Crick laid out the "Central Dogma" of molecular biology, which foretold the relationship between DNA, RNA, and proteins, and articulated the "adaptor hypothesis".^[143] Final confirmation of the replication mechanism that was implied by the double-helical structure followed in 1958 through the Meselson-Stahl experiment.^[144] Further work by Crick and coworkers showed that the genetic code was based on non-overlapping triplets of bases, called codons, allowing Har Gobind Khorana, Robert W. Holley and Marshall Warren Nirenberg to decipher the genetic code.^[145] These findings represent the birth of molecular biology.

See also

- Crystallography
- DNA microarray
- DNA sequencing
- Genetic disorder
- Junk DNA
- Molecular models of DNA
- Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid
- Nucleic acid analogues
- Nucleic acid methods
- Nucleic acid modeling
- Nucleic Acid Notations
- Paracrystal model and theory
- X-ray crystallography
- X-ray scattering
- Phosphoramidite
- Plasmid
- Polymerase chain reaction
- *Proteopedia DNA* ^[146]
- Southern blot
- Triple-stranded DNA

Further reading

- Calladine, Chris R.; Drew, Horace R.; Luisi, Ben F. and Travers, Andrew A. (2003). *Understanding DNA: the molecule & how it works*. Amsterdam: Elsevier Academic Press. ISBN 0-12-155089-3.
- Dennis, Carina; Julie Clayton (2003). *50 years of DNA*. Basingstoke: Palgrave Macmillan. ISBN 1-4039-1479-6.
- Judson, Horace Freeland (1996). *The eighth day of creation: makers of the revolution in biology*. Plainview, N.Y.: CSHL Press. ISBN 0-87969-478-5.
- Olby, Robert C. (1994). *The path to the double helix: the discovery of DNA*. New York: Dover Publications. ISBN 0-486-68117-3., first published in October 1974 by MacMillan, with foreword by Francis Crick;the definitive DNA textbook, revised in 1994 with a 9 page postscript.
- Olby, Robert C. (2009). *Francis Crick: A Biography*. Plainview, N.Y.: Cold Spring Harbor Laboratory Press. ISBN 0-87969-798-9.
- Ridley, Matt (2006). *Francis Crick: discoverer of the genetic code*. [Ashland, OH: Eminent Lives, Atlas Books. ISBN 0-06-082333-X.
- Berry, Andrew; Watson, James D. (2003). *DNA: the secret of life*. New York: Alfred A. Knopf. ISBN 0-375-41546-7.
- Stent, Gunther Siegmund; Watson, James D. (1980). *The double helix: a personal account of the discovery of the structure of DNA*. New York: Norton. ISBN 0-393-95075-1.

- Wilkins, Maurice (2003). *The third man of the double helix the autobiography of Maurice Wilkins*. Cambridge, Eng: University Press. ISBN 0-19-860665-6.

External links

- DNA ^[147] at the Open Directory Project
- DNA binding site prediction on protein ^[148]
- DNA coiling to form chromosomes ^[149]
- DNA from the Beginning ^[150] Another DNA Learning Center site on DNA, genes, and heredity from Mendel to the human genome project.
- DNA Lab, demonstrates how to extract DNA from wheat using readily available equipment and supplies. ^[151]
- DNA the Double Helix Game ^[152] From the official Nobel Prize web site
- DNA under electron microscope ^[153]
- Dolan DNA Learning Center ^[154]
- Double Helix: 50 years of DNA ^[155], *Nature*
- Double Helix 1953–2003 ^[156] National Centre for Biotechnology Education
- Francis Crick and James Watson talking on the BBC in 1962, 1972, and 1974 ^[157]
- Genetic Education Modules for Teachers ^[158]—*DNA from the Beginning* Study Guide
- Guide to DNA cloning ^[159]
- Olby R (January 2003). "Quiet debut for the double helix" ^[160]. *Nature* **421** (6921): 402–5. doi:10.1038/nature01397. PMID 12540907.
- PDB Molecule of the Month *pdb23_1* ^[161]
- Rosalind Franklin's contributions to the study of DNA ^[162]
- The Register of Francis Crick Personal Papers 1938 - 2007 ^[163] at Mandeville Special Collections Library, Geisel Library, University of California, San Diego
- U.S. National DNA Day ^[164]—watch videos and participate in real-time chat with top scientists
- "Clue to chemistry of heredity found" ^[165]. *The New York Times*. Saturday, June 13, 1953. The first American newspaper coverage of the discovery of the DNA structure.
- An Introduction to DNA and Chromosomes ^[166] from HOPES: Huntington's Disease Outreach Project for Education at Stanford

References

- [1] Russell, Peter (2001). *iGenetics*. New York: Benjamin Cummings. ISBN 0-805-34553-1.
- [2] Saenger, Wolfram (1984). *Principles of Nucleic Acid Structure*. New York: Springer-Verlag. ISBN 0387907629.
- [3] Alberts, Bruce; Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts and Peter Walters (2002). *Molecular Biology of the Cell; Fourth Edition* (<http://www.ncbi.nlm.nih.gov/books/bv.fcgi?call=bv.View..ShowTOC&rid=mboc4.TOC&depth=2>). New York and London: Garland Science. ISBN 0-8153-3218-1. OCLC 145080076 48122761 57023651 69932405. .
- [4] Butler, John M. (2001). *Forensic DNA Typing*. Elsevier. ISBN 978-0-12-147951-0. OCLC 223032110 45406517. pp. 14–15.
- [5] Mandelkern M, Elias J, Eden D, Crothers D (1981). "The dimensions of DNA in solution". *J Mol Biol* **152** (1): 153–61. doi:10.1016/0022-2836(81)90099-1. PMID 7338906.
- [6] Gregory S; Barlow, KF; McLay, KE; Kaul, R; Swarbreck, D; Dunham, A; Scott, CE; Howe, KL *et al.* (2006). "The DNA sequence and biological annotation of human chromosome 1". *Nature* **441** (7091): 315–21. doi:10.1038/nature04727. PMID 16710414.
- [7] Watson J.D. and Crick F.H.C. (1953). "A Structure for Deoxyribose Nucleic Acid" (<http://www.nature.com/nature/dna50/watsoncrick.pdf>) (PDF). *Nature* **171** (4356): 737–738. doi:10.1038/171737a0. PMID 13054692. . Retrieved 4 May 2009.
- [8] Berg J., Tymoczko J. and Stryer L. (2002) *Biochemistry*. W. H. Freeman and Company ISBN 0-7167-4955-6
- [9] Abbreviations and Symbols for Nucleic Acids, Polynucleotides and their Constituents (<http://www.chem.qmul.ac.uk/iupac/misc/naabb.html>) IUPAC-IUB Commission on Biochemical Nomenclature (CBN), Accessed 03 January 2006
- [10] Ghosh A, Bansal M (2003). "A glossary of DNA structures from A to Z". *Acta Crystallogr D Biol Crystallogr* **59** (Pt 4): 620–6. doi:10.1107/S0907444903003251. PMID 12657780.
- [11] Created from PDB 1D65 (<http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=1D65>)

- [12] Verma S, Eckstein F (1998). "Modified oligonucleotides: synthesis and strategy for users". *Annu. Rev. Biochem.* **67**: 99–134. doi:10.1146/annurev.biochem.67.1.99. PMID 9759484.
- [13] Wing R, Drew H, Takano T, Broka C, Tanaka S, Itakura K, Dickerson R (1980). "Crystal structure analysis of a complete turn of B-DNA". *Nature* **287** (5784): 755–8. doi:10.1038/287755a0. PMID 7432492.
- [14] Pabo C, Sauer R (1984). "Protein-DNA recognition". *Annu Rev Biochem* **53**: 293–321. doi:10.1146/annurev.bi.53.070184.001453. PMID 6236744.
- [15] Clausen-Schaumann H, Rief M, Tolksdorf C, Gaub H (2000). "Mechanical stability of single DNA molecules" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1300792>). *Biophys J* **78** (4): 1997–2007. doi:10.1016/S0006-3495(00)76747-6. PMID 10733978. PMC 1300792.
- [16] Yakovchuk P, Protozanova E, Frank-Kamenetskii MD (2006). "Base-stacking and base-pairing contributions into thermal stability of the DNA double helix" (<http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=16449200>). *Nucleic Acids Res.* **34** (2): 564–74. doi:10.1093/nar/gkj454. PMID 16449200. PMC 1360284. .
- [17] Chalikian T, Völker J, Plum G, Breslauer K (1999). "A more unified picture for the thermodynamics of nucleic acid duplex melting: a characterization by calorimetric and volumetric techniques" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=22151>). *Proc Natl Acad Sci USA* **96** (14): 7853–8. doi:10.1073/pnas.96.14.7853. PMID 10393911. PMC 22151.
- [18] deHaseth P, Helmann J (1995). "Open complex formation by Escherichia coli RNA polymerase: the mechanism of polymerase-induced strand separation of double helical DNA". *Mol Microbiol* **16** (5): 817–24. doi:10.1111/j.1365-2958.1995.tb02309.x. PMID 7476180.
- [19] Isaksson J, Acharya S, Barman J, Cheruku P, Chattopadhyaya J (2004). "Single-stranded adenine-rich DNA and RNA retain structural characteristics of their respective double-stranded conformations and show directional differences in stacking pattern". *Biochemistry* **43** (51): 15996–6010. doi:10.1021/bi048221v. PMID 15609994.
- [20] Designation of the two strands of DNA (<http://www.chem.qmul.ac.uk/iubmb/newsletter/misc/DNA.html>) JCBN/NC-IUB Newsletter 1989, Accessed 07 May 2008
- [21] Hüttenhofer A, Schattner P, Polacek N (2005). "Non-coding RNAs: hope or hype?". *Trends Genet* **21** (5): 289–97. doi:10.1016/j.tig.2005.03.007. PMID 15851066.
- [22] Munroe S (2004). "Diversity of antisense regulation in eukaryotes: multiple mechanisms, emerging patterns". *J Cell Biochem* **93** (4): 664–71. doi:10.1002/jcb.20252. PMID 15389973.
- [23] Makalowska I, Lin C, Makalowski W (2005). "Overlapping genes in vertebrate genomes". *Comput Biol Chem* **29** (1): 1–12. doi:10.1016/j.compbiolchem.2004.12.006. PMID 15680581.
- [24] Johnson Z, Chisholm S (2004). "Properties of overlapping genes are conserved across microbial genomes" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=525685>). *Genome Res* **14** (11): 2268–72. doi:10.1101/gr.2433104. PMID 15520290. PMC 525685.
- [25] Lamb R, Horvath C (1991). "Diversity of coding strategies in influenza viruses". *Trends Genet* **7** (8): 261–6. PMID 1771674.
- [26] Benham C, Mielke S (2005). "DNA mechanics". *Annu Rev Biomed Eng* **7**: 21–53. doi:10.1146/annurev.bioeng.6.062403.132016. PMID 16004565.
- [27] Champoux J (2001). "DNA topoisomerases: structure, function, and mechanism". *Annu Rev Biochem* **70**: 369–413. doi:10.1146/annurev.biochem.70.1.369. PMID 11395412.
- [28] Wang J (2002). "Cellular roles of DNA topoisomerases: a molecular perspective". *Nat Rev Mol Cell Biol* **3** (6): 430–40. doi:10.1038/nrm831. PMID 12042765.
- [29] Basu H, Feuerstein B, Zarling D, Shafer R, Marton L (1988). "Recognition of Z-RNA and Z-DNA determinants by polyamines in solution: experimental and theoretical studies". *J Biomol Struct Dyn* **6** (2): 299–309. PMID 2482766.
- [30] Franklin RE, Gosling RG (6 March 1953). "The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content" (http://hekto.med.unc.edu:8080/CARTER/carter_WWW/Bioch_134/PDF_files/Franklin_Gosling.pdf). *Acta Cryst* **6** (8-9): 673–7. doi:10.1107/S0365110X53001939. .
Franklin RE, Gosling RG (September 1953). "The structure of sodium thymonucleate fibres. II. The cylindrically symmetrical Patterson function". *Acta Cryst* **6** (8-9): 678–85. doi:10.1107/S0365110X53001940.
- [31] Franklin, Rosalind and Gosling, Raymond (1953). "Molecular Configuration in Sodium Thymonucleate. Franklin R. and Gosling R.G" (<http://www.nature.com/nature/dna50/franklingosling.pdf>) (PDF). *Nature* **171** (4356): 740–1. doi:10.1038/171740a0. PMID 13054694. .
- [32] Wilkins M.H.F., A.R. Stokes A.R. & Wilson, H.R. (1953). "Molecular Structure of Deoxypentose Nucleic Acids" (<http://www.nature.com/nature/dna50/wilkins.pdf>) (PDF). *Nature* **171** (4356): 738–740. doi:10.1038/171738a0. PMID 13054693. .
- [33] Leslie AG, Arnott S, Chandrasekaran R, Ratliff RL (1980). "Polymorphism of DNA double helices". *J. Mol. Biol.* **143** (1): 49–72. doi:10.1016/0022-2836(80)90124-2. PMID 7441761.
- [34] Baianu, I.C. (1980). "Structural Order and Partial Disorder in Biological systems". *Bull. Math. Biol.* **42** (4): 137–141. <http://cogprints.org/3822/>
- [35] Hosemann R., Bagchi R.N., *Direct analysis of diffraction by matter*, North-Holland Publ., Amsterdam – New York, 1962.
- [36] Baianu, I.C. (1978). "X-ray scattering by partially disordered membrane systems.". *Acta Cryst.*, **A34** (5): 751–753. doi:10.1107/S0567739478001540.
- [37] Wahl M, Sundaralingam M (1997). "Crystal structures of A-DNA duplexes". *Biopolymers* **44** (1): 45–63. doi:10.1002/(SICI)1097-0282(1997)44:1 (inactive 2009-03-14). PMID 9097733.

- [38] Lu XJ, Shakked Z, Olson WK (2000). "A-form conformational motifs in ligand-bound DNA structures". *J. Mol. Biol.* **300** (4): 819–40. doi:10.1006/jmbi.2000.3690. PMID 10891271.
- [39] Rothenburg S, Koch-Nolte F, Haag F (2001). "DNA methylation and Z-DNA formation as mediators of quantitative differences in the expression of alleles". *Immunol Rev* **184**: 286–98. doi:10.1034/j.1600-065x.2001.1840125.x. PMID 12086319.
- [40] Oh D, Kim Y, Rich A (2002). "Z-DNA-binding proteins can act as potent effectors of gene expression in vivo" (<http://www.pnas.org/cgi/pmidlookup?view=long&pmid=12486233>). *Proc. Natl. Acad. Sci. U.S.A.* **99** (26): 16666–71. doi:10.1073/pnas.262672699. PMID 12486233. PMC 139201. .
- [41] Created from NDB UD0017 (<http://ndbserver.rutgers.edu/atlas/xray/structures/U/ud0017/ud0017.html>)
- [42] Greider C, Blackburn E (1985). "Identification of a specific telomere terminal transferase activity in Tetrahymena extracts". *Cell* **43** (2 Pt 1): 405–13. doi:10.1016/0092-8674(85)90170-9. PMID 3907856.
- [43] Nugent C, Lundblad V (1998). "The telomerase reverse transcriptase: components and regulation" (<http://www.genesdev.org/cgi/content/full/12/8/1073>). *Genes Dev* **12** (8): 1073–85. doi:10.1101/gad.12.8.1073. PMID 9553037. .
- [44] Wright W, Tesmer V, Huffman K, Levene S, Shay J (1997). "Normal human chromosomes have long G-rich telomeric overhangs at one end" (<http://www.genesdev.org/cgi/content/full/11/21/2801>). *Genes Dev* **11** (21): 2801–9. doi:10.1101/gad.11.21.2801. PMID 9353250. PMC 316649. .
- [45] Burge S, Parkinson G, Hazel P, Todd A, Neidle S (2006). "Quadruplex DNA: sequence, topology and structure" (<http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=17012276>). *Nucleic Acids Res* **34** (19): 5402–15. doi:10.1093/nar/gkl655. PMID 17012276. PMC 1636468. .
- [46] Parkinson G, Lee M, Neidle S (2002). "Crystal structure of parallel quadruplexes from human telomeric DNA". *Nature* **417** (6891): 876–80. doi:10.1038/nature755. PMID 12050675.
- [47] Griffith J, Comeau L, Rosenfield S, Stansel R, Bianchi A, Moss H, de Lange T (1999). "Mammalian telomeres end in a large duplex loop". *Cell* **97** (4): 503–14. doi:10.1016/S0092-8674(00)80760-6. PMID 10338214.
- [48] Seeman NC (November 2005). "DNA enables nanoscale control of the structure of matter". *Q. Rev. Biophys.* **38** (4): 363–71. doi:10.1017/S0033583505004087. PMID 16515737.
- [49] Klose R, Bird A (2006). "Genomic DNA methylation: the mark and its mediators". *Trends Biochem Sci* **31** (2): 89–97. doi:10.1016/j.tibs.2005.12.008. PMID 16403636.
- [50] Bird A (2002). "DNA methylation patterns and epigenetic memory". *Genes Dev* **16** (1): 6–21. doi:10.1101/gad.947102. PMID 11782440.
- [51] Walsh C, Xu G (2006). "Cytosine methylation and DNA repair". *Curr Top Microbiol Immunol* **301**: 283–315. doi:10.1007/3-540-31390-7_11. PMID 16570853.
- [52] Kriaucionis S, Heintz N (May 2009). "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain". *Science* **324** (5929): 929–30. doi:10.1126/science.1169786. PMID 19372393.
- [53] Ratel D, Ravanat J, Berger F, Wion D (2006). "N6-methyladenine: the other methylated base of DNA" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2754416>). *Bioessays* **28** (3): 309–15. doi:10.1002/bies.20342. PMID 16479578. PMC 2754416.
- [54] Gommers-Ampt J, Van Leeuwen F, de Beer A, Vliegthart J, Dizdaroğlu M, Kowalak J, Crain P, Borst P (1993). "beta-D-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan *T. brucei*". *Cell* **75** (6): 1129–36. doi:10.1016/0092-8674(93)90322-H. PMID 8261512.
- [55] Created from PDB 1JDG (<http://www.rcsb.org/pdb/cgi/explore.cgi?pdbId=1JDG>)
- [56] Douki T, Reynaud-Angelin A, Cadet J, Sage E (2003). "Bipyrimidine photoproducts rather than oxidative lesions are the main type of DNA damage involved in the genotoxic effect of solar UVA radiation". *Biochemistry* **42** (30): 9221–6. doi:10.1021/bi034593c. PMID 12885257. .
- [57] Cadet J, Delatour T, Douki T, Gasparutto D, Pouget J, Ravanat J, Sauvaigo S (1999). "Hydroxyl radicals and DNA base damage". *Mutat Res* **424** (1–2): 9–21. PMID 10064846.
- [58] Beckman KB, Ames BN (August 1997). "Oxidative decay of DNA" (<http://www.jbc.org/cgi/pmidlookup?view=long&pmid=9289489>). *J. Biol. Chem.* **272** (32): 19633–6. doi:10.1074/jbc.272.32.19633. PMID 9289489. .
- [59] Valerie K, Povirk L (2003). "Regulation and mechanisms of mammalian double-strand break repair". *Oncogene* **22** (37): 5792–812. doi:10.1038/sj.onc.1206679. PMID 12947387.
- [60] Ferguson L, Denny W (1991). "The genetic toxicology of acridines". *Mutat Res* **258** (2): 123–60. PMID 1881402.
- [61] Jeffrey A (1985). "DNA modification by chemical carcinogens". *Pharmacol Ther* **28** (2): 237–72. doi:10.1016/0163-7258(85)90013-0. PMID 3936066.
- [62] Stephens T, Bunde C, Fillmore B (2000). "Mechanism of action in thalidomide teratogenesis". *Biochem Pharmacol* **59** (12): 1489–99. doi:10.1016/S0006-2952(99)00388-3. PMID 10799645.
- [63] Braña M, Cacho M, Gradillas A, de Pascual-Teresa B, Ramos A (2001). "Intercalators as anticancer drugs". *Curr Pharm Des* **7** (17): 1745–80. doi:10.2174/1381612013397113. PMID 11562309.
- [64] Venter J, Adams, MD; Myers, EW; Li, PW; Mural, RJ; Sutton, GG; Smith, HO; Yandell, M *et al.* (2001). "The sequence of the human genome". *Science* **291** (5507): 1304–51. doi:10.1126/science.1058040. PMID 11181995.
- [65] Thanbichler M, Wang S, Shapiro L (2005). "The bacterial nucleoid: a highly organized and dynamic structure". *J Cell Biochem* **96** (3): 506–21. doi:10.1002/jcb.20519. PMID 15988757.
- [66] Wolfsberg T, McEntyre J, Schuler G (2001). "Guide to the draft human genome". *Nature* **409** (6822): 824–6. doi:10.1038/35057000. PMID 11236998.

- [67] Gregory T (2005). "The C-value enigma in plants and animals: a review of parallels and an appeal for partnership" (<http://aob.oxfordjournals.org/cgi/content/full/95/1/133>). *Ann Bot (Lond)* **95** (1): 133–46. doi:10.1093/aob/mci009. PMID 15596463. .
- [68] The ENCODE Project Consortium (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=2212820>). *Nature* **447** (7146): 799–816. doi:10.1038/nature05874. PMID 17571346. PMC 2212820.
- [69] Created from PDB 1MSW (<http://www.rcsb.org/pdb/explore/explore.do?structureId=1MSW>)
- [70] Pidoux A, Allshire R (2005). "The role of heterochromatin in centromere function" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=1569473>). *Philos Trans R Soc Lond B Biol Sci* **360** (1455): 569–79. doi:10.1098/rstb.2004.1611. PMID 15905142. PMC 1569473.
- [71] Harrison P, Hegyi H, Balasubramanian S, Luscombe N, Bertone P, Echols N, Johnson T, Gerstein M (2002). "Molecular fossils in the human genome: identification and analysis of the pseudogenes in chromosomes 21 and 22" (<http://www.genome.org/cgi/content/full/12/2/272>). *Genome Res* **12** (2): 272–80. doi:10.1101/gr.207102. PMID 11827946. PMC 155275. .
- [72] Harrison P, Gerstein M (2002). "Studying genomes through the aeons: protein families, pseudogenes and proteome evolution". *J Mol Biol* **318** (5): 1155–74. doi:10.1016/S0022-2836(02)00109-2. PMID 12083509.
- [73] Albà M (2001). "Replicative DNA polymerases" (<http://genomebiology.com/1465-6906/2/REVIEWS3002>). *Genome Biol* **2** (1): REVIEWS3002. doi:10.1186/gb-2001-2-1-reviews3002. PMID 11178285. PMC 150442. .
- [74] Sandman K, Pereira S, Reeve J (1998). "Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome". *Cell Mol Life Sci* **54** (12): 1350–64. doi:10.1007/s000180050259. PMID 9893710.
- [75] Dame RT (2005). "The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin". *Mol. Microbiol.* **56** (4): 858–70. doi:10.1111/j.1365-2958.2005.04598.x. PMID 15853876.
- [76] Luger K, Mäder A, Richmond R, Sargent D, Richmond T (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution". *Nature* **389** (6648): 251–60. doi:10.1038/38444. PMID 9305837.
- [77] Jenuwein T, Allis C (2001). "Translating the histone code". *Science* **293** (5532): 1074–80. doi:10.1126/science.1063127. PMID 11498575.
- [78] Ito T (2003). "Nucleosome assembly and remodelling". *Curr Top Microbiol Immunol* **274**: 1–22. PMID 12596902.
- [79] Thomas J (2001). "HMG1 and 2: architectural DNA-binding proteins". *Biochem Soc Trans* **29** (Pt 4): 395–401. doi:10.1042/BST0290395. PMID 11497996.
- [80] Grosschedl R, Giese K, Pagel J (1994). "HMG domain proteins: architectural elements in the assembly of nucleoprotein structures". *Trends Genet* **10** (3): 94–100. doi:10.1016/0168-9525(94)90232-1. PMID 8178371.
- [81] Iftode C, Daniely Y, Borowiec J (1999). "Replication protein A (RPA): the eukaryotic SSB". *Crit Rev Biochem Mol Biol* **34** (3): 141–80. doi:10.1080/10409239991209255. PMID 10473346.
- [82] Created from PDB 1LMB (<http://www.rcsb.org/pdb/explore/explore.do?structureId=1LMB>)
- [83] Myers L, Kornberg R (2000). "Mediator of transcriptional regulation". *Annu Rev Biochem* **69**: 729–49. doi:10.1146/annurev.biochem.69.1.729. PMID 10966474.
- [84] Spiegelman B, Heinrich R (2004). "Biological control through regulated transcriptional coactivators". *Cell* **119** (2): 157–67. doi:10.1016/j.cell.2004.09.037. PMID 15479634.
- [85] Li Z, Van Calcar S, Qu C, Cavenne W, Zhang M, Ren B (2003). "A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells" (<http://www.pnas.org/cgi/pmidlookup?view=long&pmid=12808131>). *Proc Natl Acad Sci USA* **100** (14): 8164–9. doi:10.1073/pnas.1332764100. PMID 12808131. PMC 166200. .
- [86] Pabo C, Sauer R (1984). "Protein-DNA recognition". *Annu Rev Biochem* **53**: 293–321. doi:10.1146/annurev.bi.53.070184.001453. PMID 6236744.
- [87] Created from PDB 1RVA (<http://www.rcsb.org/pdb/explore/explore.do?structureId=1RVA>)
- [88] Bickle T, Krüger D (1993). "Biology of DNA restriction" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=372918>). *Microbiol Rev* **57** (2): 434–50. PMID 8336674. PMC 372918.
- [89] Doherty A, Suh S (2000). "Structural and mechanistic conservation in DNA ligases" (<http://nar.oxfordjournals.org/cgi/pmidlookup?view=long&pmid=11058099>). *Nucleic Acids Res* **28** (21): 4051–8. doi:10.1093/nar/28.21.4051. PMID 11058099. PMC 113121. .
- [90] Schoeffler A, Berger J (2005). "Recent advances in understanding structure-function relationships in the type II topoisomerase mechanism". *Biochem Soc Trans* **33** (Pt 6): 1465–70. doi:10.1042/BST20051465. PMID 16246147.
- [91] Tuteja N, Tuteja R (2004). "Unraveling DNA helicases. Motif, structure, mechanism and function". *Eur J Biochem* **271** (10): 1849–63. doi:10.1111/j.1432-1033.2004.04094.x. PMID 15128295.
- [92] Joyce C, Steitz T (1995). "Polymerase structures and function: variations on a theme?" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=177480>). *J Bacteriol* **177** (22): 6321–9. PMID 7592405. PMC 177480.
- [93] Hubscher U, Maga G, Spadari S (2002). "Eukaryotic DNA polymerases". *Annu Rev Biochem* **71**: 133–63. doi:10.1146/annurev.biochem.71.090501.150041. PMID 12045093.
- [94] Johnson A, O'Donnell M (2005). "Cellular DNA replicases: components and dynamics at the replication fork". *Annu Rev Biochem* **74**: 283–315. doi:10.1146/annurev.biochem.73.011303.073859. PMID 15952889.
- [95] Tarrago-Litvak L, Andréola M, Nevinsky G, Sarih-Cottin L, Litvak S (1 May 1994). "The reverse transcriptase of HIV-1: from enzymology to therapeutic intervention" (<http://www.fasebj.org/cgi/reprint/8/8/497>). *FASEB J* **8** (8): 497–503. PMID 7514143. .

- [96] Martinez E (2002). "Multi-protein complexes in eukaryotic gene transcription". *Plant Mol Biol* **50** (6): 925–47. doi:10.1023/A:1021258713850. PMID 12516863.
- [97] Created from PDB 1M6G (<http://www.rcsb.org/pdb/explore/explore.do?structureId=1M6G>)
- [98] Cremer T, Cremer C (2001). "Chromosome territories, nuclear architecture and gene regulation in mammalian cells". *Nat Rev Genet* **2** (4): 292–301. doi:10.1038/35066075. PMID 11283701.
- [99] Pál C, Papp B, Lercher M (2006). "An integrated view of protein evolution". *Nat Rev Genet* **7** (5): 337–48. doi:10.1038/nrg1838. PMID 16619049.
- [100] O'Driscoll M, Jeggo P (2006). "The role of double-strand break repair - insights from human genetics". *Nat Rev Genet* **7** (1): 45–54. doi:10.1038/nrg1746. PMID 16369571.
- [101] Vispé S, Defais M (1997). "Mammalian Rad51 protein: a RecA homologue with pleiotropic functions". *Biochimie* **79** (9-10): 587–92. doi:10.1016/S0300-9084(97)82007-X. PMID 9466696.
- [102] Neale MJ, Keeney S (2006). "Clarifying the mechanics of DNA strand exchange in meiotic recombination". *Nature* **442** (7099): 153–8. doi:10.1038/nature04885. PMID 16838012.
- [103] Dickman M, Ingleston S, Sedelnikova S, Rafferty J, Lloyd R, Grasby J, Hornby D (2002). "The RuvABC resolvosome". *Eur J Biochem* **269** (22): 5492–501. doi:10.1046/j.1432-1033.2002.03250.x. PMID 12423347.
- [104] Orgel L (2004). "Prebiotic chemistry and the origin of the RNA world" (<http://www.crbmb.com/cgi/reprint/39/2/99.pdf>) (PDF). *Crit Rev Biochem Mol Biol* **39** (2): 99–123. doi:10.1080/10409230490460765. PMID 15217990. .
- [105] Davenport R (2001). "Ribozymes. Making copies in the RNA world". *Science* **292** (5520): 1278. doi:10.1126/science.292.5520.1278a. PMID 11360970.
- [106] Szathmáry E (1992). "What is the optimum size for the genetic alphabet?" (<http://www.pnas.org/cgi/reprint/89/7/2614.pdf>) (PDF). *Proc Natl Acad Sci USA* **89** (7): 2614–8. doi:10.1073/pnas.89.7.2614. PMID 1372984. PMC 48712. .
- [107] Lindahl T (1993). "Instability and decay of the primary structure of DNA". *Nature* **362** (6422): 709–15. doi:10.1038/362709a0. PMID 8469282.
- [108] Vreeland R, Rosenzweig W, Powers D (2000). "Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal". *Nature* **407** (6806): 897–900. doi:10.1038/35038060. PMID 11057666.
- [109] Hebsgaard M, Phillips M, Willerslev E (2005). "Geologically ancient DNA: fact or artefact?". *Trends Microbiol* **13** (5): 212–20. doi:10.1016/j.tim.2005.03.010. PMID 15866038.
- [110] Nickle D, Learn G, Rain M, Mullins J, Mittler J (2002). "Curiously modern DNA for a "250 million-year-old" bacterium". *J Mol Evol* **54** (1): 134–7. doi:10.1007/s00239-001-0025-x. PMID 11734907.
- [111] Goff SP, Berg P (1976). "Construction of hybrid viruses containing SV40 and lambda phage DNA segments and their propagation in cultured monkey cells". *Cell* **9** (4 PT 2): 695–705. doi:10.1016/0092-8674(76)90133-1. PMID 189942.
- [112] Houdebine L (2007). "Transgenic animal models in biomedical research". *Methods Mol Biol* **360**: 163–202. doi:10.1385/1-59745-165-7:163. PMID 17172731.
- [113] Daniell H, Dhingra A (2002). "Multigene engineering: dawn of an exciting new era in biotechnology". *Curr Opin Biotechnol* **13** (2): 136–41. doi:10.1016/S0958-1669(02)00297-5. PMID 11950565.
- [114] Job D (2002). "Plant biotechnology in agriculture". *Biochimie* **84** (11): 1105–10. doi:10.1016/S0300-9084(02)00013-5. PMID 12595138.
- [115] Collins A, Morton N (1994). "Likelihood ratios for DNA identification" (<http://www.pnas.org/cgi/reprint/91/13/6007.pdf>) (PDF). *Proc Natl Acad Sci USA* **91** (13): 6007–11. doi:10.1073/pnas.91.13.6007. PMID 8016106. PMC 44126. .
- [116] Weir B, Triggs C, Starling L, Stowell L, Walsh K, Buckleton J (1997). "Interpreting DNA mixtures". *J Forensic Sci* **42** (2): 213–22. PMID 9068179.
- [117] Jeffreys A, Wilson V, Thein S (1985). "Individual-specific 'fingerprints' of human DNA". *Nature* **316** (6023): 76–9. doi:10.1038/316076a0. PMID 2989708.
- [118] Colin Pitchfork — first murder conviction on DNA evidence also clears the prime suspect (http://www.forensic.gov.uk/forensic_t/inside/news/list_casefiles.php?case=1) Forensic Science Service Accessed 23 December 2006
- [119] "DNA Identification in Mass Fatality Incidents" (<http://massfatality.dna.gov/Introduction/>). National Institute of Justice. September 2006. .
- [120] Baldi, Pierre; Brunak, Soren (2001). *Bioinformatics: The Machine Learning Approach*. MIT Press. ISBN 978-0-262-02506-5. OCLC 45951728..
- [121] Gusfield, Dan. *Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology*. Cambridge University Press, 15 January 1997. ISBN 978-0-521-58519-4.
- [122] Sjölander K (2004). "Phylogenomic inference of protein molecular function: advances and challenges" (<http://bioinformatics.oxfordjournals.org/cgi/reprint/20/2/170>). *Bioinformatics* **20** (2): 170–9. doi:10.1093/bioinformatics/bth021. PMID 14734307. .
- [123] Mount DM (2004). *Bioinformatics: Sequence and Genome Analysis* (2 ed.). Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. ISBN 0879697121. OCLC 55106399.
- [124] Rothmund PW (March 2006). "Folding DNA to create nanoscale shapes and patterns". *Nature* **440** (7082): 297–302. doi:10.1038/nature04586. PMID 16541064.
- [125] Andersen ES, Dong M, Nielsen MM (May 2009). "Self-assembly of a nanoscale DNA box with a controllable lid". *Nature* **459** (7243): 73–6. doi:10.1038/nature07971. PMID 19424153.

- [126] Ishitsuka Y, Ha T (May 2009). "DNA nanotechnology: a nanomachine goes live". *Nat Nanotechnol* **4** (5): 281–2. doi:10.1038/nnano.2009.101. PMID 19421208.
- [127] Aldaye FA, Palmer AL, Sleiman HF (September 2008). "Assembling materials with DNA as the guide". *Science* **321** (5897): 1795–9. doi:10.1126/science.1154533. PMID 18818351.
- [128] Wray G; Martindale, Mark Q. (2002). "Dating branches on the tree of life using DNA" (<http://genomebiology.com/1465-6906/3/REVIEWS0001>). *Genome Biol* **3** (1): REVIEWS0001. doi:10.1046/j.1525-142X.1999.99010.x. PMID 11806830. PMC 150454. .
- [129] *Lost Tribes of Israel*, NOVA, PBS airdate: 22 February 2000. Transcript available from PBS.org, (<http://www.pbs.org/wgbh/nova/transcripts/2706israel.html>) (last accessed on 4 March 2006)
- [130] Kleiman, Yaakov. "The Cohanim/DNA Connection: The fascinating story of how DNA studies confirm an ancient biblical tradition". (http://www.aish.com/societywork/sciencenature/the_cohanim_-_dna_connection.asp) *aish.com* (January 13, 2000). Accessed 4 March 2006.
- [131] Bhattacharya, Shaoni. "Killer convicted thanks to relative's DNA". (<http://www.newscientist.com/article.ns?id=dn4908>) *newscientist.com* (20 April 2004). Accessed 22 December 06
- [132] Dahm R (January 2008). "Discovering DNA: Friedrich Miescher and the early years of nucleic acid research". *Hum. Genet.* **122** (6): 565–81. doi:10.1007/s00439-007-0433-0. PMID 17901982.
- [133] Levene P, (1 December 1919). "The structure of yeast nucleic acid" (<http://www.jbc.org/cgi/reprint/40/2/415>). *J Biol Chem* **40** (2): 415–24. .
- [134] Astbury W, (1947). "Nucleic acid". *Symp. SOC. Exp. Bbl* **1** (66).
- [135] Lorenz MG, Wackernagel W (1 September 1994). "Bacterial gene transfer by natural genetic transformation in the environment" (<http://mmbr.asm.org/cgi/pmidlookup?view=long&pmid=7968924>). *Microbiol. Rev.* **58** (3): 563–602. PMID 7968924. PMC 372978. .
- [136] Avery O, MacLeod C, McCarty M (1944). "Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III" (<http://www.jem.org/cgi/reprint/149/2/297>). *J Exp Med* **79** (2): 137–158. doi:10.1084/jem.79.2.137. PMID 19871359. PMC 2135445. .
- [137] Hershey A, Chase M (1952). "Independent functions of viral protein and nucleic acid in growth of bacteriophage" (<http://www.jgp.org/cgi/reprint/36/1/39.pdf>) (PDF). *J Gen Physiol* **36** (1): 39–56. doi:10.1085/jgp.36.1.39. PMID 12981234. PMC 2147348. .
- [138] The B-DNA X-ray pattern on the right of this linked image (<http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/dna/pictures/sci9.001.5.html>) was obtained by Rosalind Franklin and Raymond Gosling in May 1952 at high hydration levels of DNA and it has been labeled as "Photo 51"
- [139] Nature Archives Double Helix of DNA: 50 Years (<http://www.nature.com/nature/dna50/archive.html>)
- [140] Original X-ray diffraction image (<http://osulibrary.oregonstate.edu/specialcollections/coll/pauling/dna/pictures/franklin-typeBphoto.html>)
- [141] The Nobel Prize in Physiology or Medicine 1962 (http://nobelprize.org/nobel_prizes/medicine/laureates/1962/) Nobelprize .org Accessed 22 December 06
- [142] Brenda Maddox (23 January 2003). "The double helix and the 'wronged heroine'" (http://www.biomath.nyu.edu/index/course/hw_articles/nature4.pdf) (PDF). *Nature* **421** (6921): 407–408. doi:10.1038/nature01399. PMID 12540909. .
- [143] Crick, F.H.C. On degenerate templates and the adaptor hypothesis (PDF). (<http://genome.wellcome.ac.uk/assets/wtx030893.pdf>) genome.wellcome.ac.uk (Lecture, 1955). Accessed 22 December 2006
- [144] Meselson M, Stahl F (1958). "The replication of DNA in *Escherichia coli*" (<http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pmcentrez&artid=528642>). *Proc Natl Acad Sci USA* **44** (7): 671–82. doi:10.1073/pnas.44.7.671. PMID 16590258. PMC 528642.
- [145] The Nobel Prize in Physiology or Medicine 1968 (http://nobelprize.org/nobel_prizes/medicine/laureates/1968/) Nobelprize.org Accessed 22 December 06
- [146] <http://proteopedia.org/wiki/index.php/DNA>
- [147] http://www.dmoz.org/Science/Biology/Biochemistry_and_Molecular_Biology/Biomolecules/Nucleic_Acids/DNA/
- [148] <http://pipe.scs.fsu.edu/displar.html>
- [149] http://biostudio.com/c_%20education%20mac.htm
- [150] <http://www.dnaftb.org/dnaftb/>
- [151] <http://ca.youtube.com/watch?v=iyb7fwduuGM>
- [152] http://nobelprize.org/educational_games/medicine/dna_double_helix/
- [153] http://www.fidelitysystems.com/Unlinked_DNA.html
- [154] <http://www.dnalc.org/>
- [155] <http://www.nature.com/nature/dna50/archive.html>
- [156] <http://www.ncbe.reading.ac.uk/DNA50/>
- [157] <http://www.bbc.co.uk/bbcfour/audiointerviews/profilepages/crickwatson1.shtml>
- [158] <http://www.genome.gov/10506718>
- [159] <http://www.blackwellpublishing.com/trun/artwork/Animations/cloningexp/cloningexp.html>
- [160] <http://chem-faculty.ucsd.edu/joseph/CHEM13/DNA1.pdf>
- [161] http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb23_1.html
- [162] <http://mason.gmu.edu/~emoody/rfranklin.html>
- [163] <http://orpheus.ucsd.edu/speccoll/testing/html/mss0660a.html#abstract>

[164] <http://www.genome.gov/10506367>

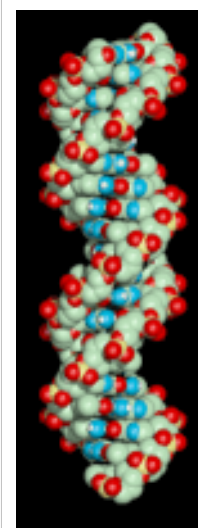
[165] <http://www.nytimes.com/packages/pdf/science/dna-article.pdf>

[166] <http://hopes.stanford.edu/basics/dna/b0.html>

Molecular models of DNA

Molecular models of DNA structures are representations of the molecular geometry and topology of Deoxyribonucleic acid (DNA) molecules using one of several means, with the aim of simplifying and presenting the essential, physical and chemical, properties of DNA molecular structures either *in vivo* or *in vitro*. These representations include closely packed spheres (CPK models) made of plastic, metal wires for 'skeletal models', graphic computations and animations by computers, artistic rendering. Computer molecular models also allow animations and molecular dynamics simulations that are very important for understanding how DNA functions *in vivo*.

The more advanced, computer-based molecular models of DNA involve molecular dynamics simulations as well as quantum mechanical computations of vibro-rotations, delocalized molecular orbitals (MOs), electric dipole moments, hydrogen-bonding, and so on. **DNA molecular dynamics modeling** involves simulations of DNA molecular geometry and topology changes with time as a result of both intra- and inter- molecular interactions of DNA. Whereas molecular models of Deoxyribonucleic acid (DNA) molecules such as closely packed spheres (CPK models) made of plastic or metal wires for 'skeletal models' are useful representations of static DNA structures, their usefulness is very limited for representing complex DNA dynamics. Computer molecular modeling allows both animations and molecular dynamics simulations that are very important for understanding how DNA functions *in vivo*.



Spinning DNA
generic model.

History

Double Helix Discovery



William Astbury

Oswald Avery

Francis Crick

Erwin Chargaff

Max Delbrück

Jerry Donohue

Rosalind Franklin

Raymond Gosling

Phoebus Levene

Linus Pauling

Sir John Randall

Erwin Schrödinger

Alex Stokes

James Watson

Maurice Wilkins

Herbert Wilson

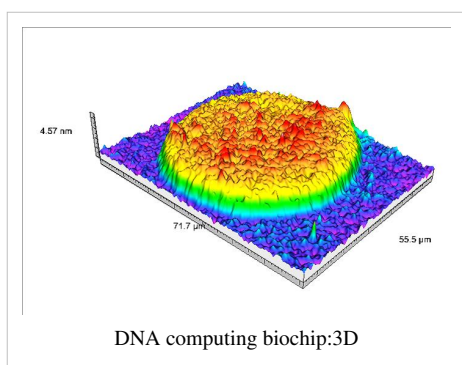
From the very early stages of structural studies of DNA by X-ray diffraction and biochemical means, molecular models such as the Watson-Crick double-helix model were successfully employed to solve the 'puzzle' of DNA structure, and also find how the latter relates to its key functions in living cells. The first high quality X-ray diffraction patterns of A-DNA were reported by Rosalind Franklin and Raymond Gosling in 1953^[1]. The first calculations of the Fourier transform of an atomic helix were reported one year earlier by Cochran, Crick and Vand^[2], and were followed in 1953 by the computation of the Fourier transform of a coiled-coil by Crick^[3].

Structural information is generated from X-ray diffraction studies of oriented DNA fibers with the help of molecular models of DNA that are combined with crystallographic and mathematical analysis of the X-ray patterns.

The first reports of a double-helix molecular model of B-DNA structure were made by Watson and Crick in 1953^[4]^[5]. Last-but-not-least, Maurice F. Wilkins, A. Stokes and H.R. Wilson, reported the first X-ray patterns of *in vivo* B-DNA in partially oriented salmon sperm heads^[6]. The development of the first correct double-helix molecular model of DNA by Crick and Watson may not have been possible without the biochemical evidence for the nucleotide base-pairing ([A---T]; [C---G]), or Chargaff's rules^[7] [8] [9] [10] [11] [12]. Although such initial studies of DNA structures with the help of molecular models were essentially static, their consequences for explaining the *in vivo* functions of DNA were significant in the areas of protein biosynthesis and the quasi-universality of the genetic code. Epigenetic transformation studies of DNA *in vivo* were however much slower to develop in spite of their importance for embryology, morphogenesis and cancer research. Such chemical dynamics and biochemical reactions of DNA are much more complex than the molecular dynamics of DNA physical interactions with water, ions and proteins/enzymes in living cells.

Importance

An old standing dynamic problem is how DNA "self-replication" takes place in living cells that should involve transient uncoiling of supercoiled DNA fibers. Although DNA consists of relatively rigid, very large elongated biopolymer molecules called "fibers" or chains (that are made of repeating nucleotide units of four basic types, attached to deoxyribose and phosphate groups), its molecular structure *in vivo* undergoes dynamic configuration changes that involve dynamically attached water molecules and ions. Supercoiling, packing with histones in chromosome structures, and other such supramolecular aspects also involve *in vivo* DNA topology which is even more complex than DNA molecular geometry, thus turning molecular modeling of DNA into an especially challenging problem for both molecular biologists and biotechnologists. Like other large molecules and biopolymers, DNA often exists in multiple stable geometries (that is, it exhibits conformational isomerism) and configurational, quantum states which are close to each other in energy on the potential energy surface of the DNA molecule.



Such varying molecular geometries can also be computed, at least in principle, by employing *ab initio* quantum chemistry methods that can attain high accuracy for small molecules, although claims that acceptable accuracy can be also achieved for polynucleotides, as well as DNA conformations, were recently made on the basis of VCD spectral data. Such quantum geometries define an important class of *ab initio* molecular models of DNA whose exploration has barely started especially in connection with results obtained by VCD in solutions. More detailed comparisons with such *ab initio* quantum computations are in principle obtainable through 2D-FT NMR spectroscopy and relaxation studies of polynucleotide solutions or specifically labeled DNA, as for example with deuterium labels.

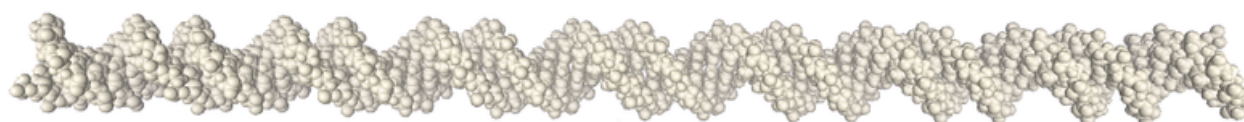
In an interesting twist of roles, the DNA molecule itself was proposed to be utilized for quantum computing. Both DNA nanostructures as well as DNA 'computing' biochips have been built (see biochip image at left).

Examples of DNA molecular models

Animated molecular models allow one to visually explore the three-dimensional (3D) structure of DNA. One visualization of DNA model is a space-filling, or CPK, model. Another is a wire, or skeletal type.

The hydrogen bonding dynamics and proton exchange is very different by many orders of magnitude between the two systems of fully hydrated DNA and water molecules in ice. Thus, the DNA dynamics is complex, involving nanosecond and several tens of picosecond time scales, whereas that of liquid ice is on the picosecond time scale, and that of proton exchange in ice is on the millisecond time scale; the proton exchange rates in DNA and attached proteins may vary from picosecond to nanosecond, minutes or years, depending on the exact locations of the exchanged protons in the large biopolymers.

A simple harmonic oscillator 'vibration' is only an oversimplified dynamic representation of the longitudinal vibrations of the DNA intertwined helices which were found to be anharmonic rather than harmonic as often assumed in quantum dynamic simulations of DNA.



DNA Spacefilling molecular model

Paracrystalline lattice models of B-DNA structures

A paracrystalline lattice, or paracrystal, is a molecular or atomic lattice with significant amounts (e.g., larger than a few percent) of partial disordering of molecular arrangements. Limiting cases of the paracrystal model are nanostructures, such as glasses, liquids, etc., that may possess only local ordering and no global order. Liquid crystals also have paracrystalline rather than crystalline structures.

Highly hydrated B-DNA occurs naturally in living cells in such a paracrystalline state, which is a dynamic one in spite of the relatively rigid DNA double-helix stabilized by parallel hydrogen bonds between the nucleotide base-pairs in the two complementary, helical DNA chains (see figures). For simplicity most DNA molecular models omit both water and ions dynamically bound to B-DNA, and are thus less useful for understanding the dynamic behaviors of B-DNA *in vivo*. The physical and mathematical analysis of X-ray^[13] ^[14] and spectroscopic data for paracrystalline B-DNA is therefore much more complicated than that of crystalline, A-DNA X-ray diffraction patterns. The paracrystal model is also important for DNA technological applications such as DNA nanotechnology.

Novel techniques that combine X-ray diffraction of DNA with X-ray microscopy in hydrated living cells are now also being developed (see, for example, "Application of X-ray microscopy in the analysis of living hydrated cells" [15]).

Genomic and biotechnology applications of DNA molecular modeling

There are various uses of DNA molecular modeling in Genomics and Biotechnology research applications, from DNA repair to PCR and DNA nanostructures. Two-dimensional DNA junction arrays have been visualized by Atomic force microscopy.^[16]

Quadruplex DNA^[17] may be involved in certain cancers^{[18] [19]}.

See also

- DNA structure
- X-ray scattering
- Neutron scattering
- Crystallography
- Crystal lattices
- 2D-FT NMRI and Spectroscopy
- List of nucleic acid simulation software
- Sirius visualization software
- X-ray microscopy
- QMC@Home
- Sir Lawrence Bragg, FRS

Spectroscopy

- Vibrational circular dichroism (VCD)
- FT-NMR^{[20] [21]}
 - NMR Atlas--database^[22]
 - mmCIF downloadable coordinate files of nucleic acids in solution from 2D-FT NMR data^[23]
 - NMR constraints files for NAs in PDB format^[24]
- NMR microscopy^[25]
- Microwave spectroscopy
- FT-IR
- FT-NIR^{[26] [27] [28]}
- Spectral, Hyperspectral, and Chemical imaging)^{[29] [30] [31] [26] [27] [32] [33]}.
- Raman spectroscopy/microscopy^[34] and CARS^[35].
- Fluorescence correlation spectroscopy^{[36] [37] [38] [27] [39] [40] [41] [26]}, Fluorescence cross-correlation spectroscopy and FRET^{[42] [43] [44]}.
- Confocal microscopy^[36]

Further reading

- *Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology*. (June 2004) I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin., q-bio/0406047.
- F. Bessel, *Untersuchung des Theils der planetarischen Störungen*, Berlin Abhandlungen (1824), article 14.
- Sir Lawrence Bragg, FRS. *The Crystalline State, A General survey*. London: G. Bells and Sons, Ltd., vols. 1 and 2., 1966., 2024 pages.
- Cantor, C. R. and Schimmel, P.R. *Biophysical Chemistry, Parts I and II.*, San Franscisco: W.H. Freeman and Co. 1980. 1,800 pages.
- Voet, D. and J.G. Voet. *Biochemistry*, 2nd Edn., New York, Toronto, Singapore: John Wiley & Sons, Inc., 1995, ISBN 0-471-58651-X., 1361 pages.
- Watson, G. N. *A Treatise on the Theory of Bessel Functions.*, (1995) Cambridge University Press. ISBN 0-521-48391-3.
- Watson, James D. *Molecular Biology of the Gene*. New York and Amsterdam: W.A. Benjamin, Inc. 1965., 494 pages.
- Wentworth, W.E. *Physical Chemistry. A short course.*, Malden (Mass.): Blackwell Science, Inc. 2000.
- Herbert R. Wilson, FRS. *Diffraction of X-rays by proteins, Nucleic Acids and Viruses.*, London: Edward Arnold (Publishers) Ltd. 1966.
- Kurt Wuthrich. *NMR of Proteins and Nucleic Acids.*, New York, Brisbane,Chicester, Toronto, Singapore: J. Wiley & Sons. 1986., 292 pages.
- Robinson, Bruche H.; Seeman, Nadrian C. (August 1987). "The Design of a Biochip: A Self-Assembling Molecular-Scale Memory Device". *Protein Engineering* **1** (4): 295–300. ISSN 0269-2139. Link ^[45]
- Rothmund, Paul W. K.; Ekani-Nkodo, Axel; Papadakis, Nick; Kumar, Ashish; Fygenon, Deborah Kuchnir & Winfree, Erik (22 December 2004). "Design and Characterization of Programmable DNA Nanotubes". *Journal of the American Chemical Society* **126** (50): 16344–16352. doi:10.1021/ja044319l. ISSN 0002-7863.
- Keren, K.; Kinneret Keren, Rotem S. Berman, Evgeny Buchstab, Uri Sivan, Erez Braun (November 2003). "DNA-Templated Carbon Nanotube Field-Effect Transistor" ^[46]. *Science* **302** (6549): 1380–1382. doi:10.1126/science.1091022. ISSN 1095-9203.
- Zheng, Jiwen; Constantinou, Pamela E.; Micheel, Christine; Alivisatos, A. Paul; Kiehl, Richard A. & Seeman Nadrian C. (2006). "2D Nanoparticle Arrays Show the Organizational Power of Robust DNA Motifs". *Nano Letters* **6**: 1502–1504. doi:10.1021/nl060994c. ISSN 1530-6984.
- Cohen, Justin D.; Sadowski, John P.; Dervan, Peter B. (2007). "Addressing Single Molecules on DNA Nanostructures". *Angewandte Chemie* **46** (42): 7956–7959. doi:10.1002/anie.200702767. ISSN 0570-0833.
- Constantinou, Pamela E.; Wang, Tong; Kopatsch, Jens; Israel, Lisa B.; Zhang, Xiaoping; Ding, Baoquan; Sherman, William B.; Wang, Xing; Zheng, Jianping; Sha, Ruojie & Seeman, Nadrian C. (2006). "Double cohesion in structural DNA nanotechnology". *Organic and Biomolecular Chemistry* **4**: 3414–3419. doi:10.1039/b605212f.
- Hallin PF, David Ussery D (2004). "CBS Genome Atlas Database: A dynamic storage for bioinformatic results and DNA sequence data". *Bioinformatics* **20**: 3682–3686.
- Zhang CT, Zhang R, Ou HY (2003). "The Z curve database: a graphic representation of genome sequences". *Bioinformatics* **19** (5): 593-599. doi:10.1093/bioinformatics/btg041

External links

- DNA the Double Helix Game ^[152] From the official Nobel Prize web site
- MDDNA: Structural Bioinformatics of DNA ^[47]
- Double Helix 1953–2003 ^[156] National Centre for Biotechnology Education
- DNALive: a web interface to compute DNA physical properties ^[48]. Also allows cross-linking of the results with the UCSC Genome browser and DNA dynamics.
- DiProDB: Dinucleotide Property Database ^[49]. The database is designed to collect and analyse thermodynamic, structural and other dinucleotide properties.
- Further details of mathematical and molecular analysis of DNA structure based on X-ray data ^[50]
- Bessel functions corresponding to Fourier transforms of atomic or molecular helices. ^[51]
- Application of X-ray microscopy in analysis of living hydrated cells ^[15]
- overview of STM/AFM/SNOM principles with educative videos ^[52]

Databases for DNA molecular models and sequences

X-ray diffraction

- NDB ID: UD0017 Database ^[17]
- X-ray Atlas -database ^[53]
- PDB files of coordinates for nucleic acid structures from X-ray diffraction by NA (incl. DNA) crystals ^[54]
- Structure factors downloadable files in CIF format ^[55]

Neutron scattering

- ISIS neutron source: ISIS pulsed neutron source: A world centre for science with neutrons & muons at Harwell, near Oxford, UK. ^[56]

X-ray microscopy

- Application of X-ray microscopy in the analysis of living hydrated cells ^[15]

Electron microscopy

- DNA under electron microscope ^[153]

Genomic and structural databases

- CBS Genome Atlas Database ^[57] — contains examples of base skews.
- The Z curve database of genomes — a 3-dimensional visualization and analysis tool of genomes ^[58].
- DNA and other nucleic acids' molecular models: Coordinate files of nucleic acids molecular structure models in PDB and CIF formats ^[59]

Atomic force microscopy

- How SPM Works ^[60]
 - SPM Image Gallery - AFM STM SEM MFM NSOM and more. ^[61]
-

References

- [1] Franklin, R.E. and Gosling, R.G. recd.6 March 1953. *Acta Cryst.* (1953). 6, 673 The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content *Acta Cryst.* (1953). and 6, 678 The Structure of Sodium Thymonucleate Fibres II. The Cylindrically Symmetrical Patterson Function.
- [2] Cochran, W., Crick, F.H.C. and Vand V. 1952. The Structure of Synthetic Polypeptides. 1. The Transform of Atoms on a Helix. *Acta Cryst.* 5(5):581-586.
- [3] Crick, F.H.C. 1953a. The Fourier Transform of a Coiled-Coil., *Acta Crystallographica* 6(8-9):685-689.
- [4] Watson, James D. and Francis H.C. Crick. A structure for Deoxyribose Nucleic Acid (<http://www.nature.com/nature/dna50/watsoncrick.pdf>) (PDF). *Nature* 171, 737–738, 25 April 1953.
- [5] Watson, J.D; Crick F.H.C. 1953b. The Structure of DNA., Cold Spring Harbor Symposia on Quantitative Biology 18:123-131.
- [6] Wilkins M.H.F., A.R. Stokes A.R. & Wilson, H.R. (1953). "Molecular Structure of Deoxypentose Nucleic Acids" (<http://www.nature.com/nature/dna50/wilkins.pdf>) (PDF). *Nature* **171** (4356): 738–740. doi:10.1038/171738a0. PMID 13054693. .
- [7] Elson D, Chargaff E (1952). "On the deoxyribonucleic acid content of sea urchin gametes". *Experientia* **8** (4): 143–145.
- [8] Chargaff E, Lipshitz R, Green C (1952). "Composition of the deoxypentose nucleic acids of four genera of sea-urchin". *J Biol Chem* **195** (1): 155–160. PMID 14938364.
- [9] Chargaff E, Lipshitz R, Green C, Hodes ME (1951). "The composition of the deoxyribonucleic acid of salmon sperm". *J Biol Chem* **192** (1): 223–230. PMID 14917668.
- [10] Chargaff E (1951). "Some recent studies on the composition and structure of nucleic acids". *J Cell Physiol Suppl* **38** (Suppl).
- [11] Magasanik B, Vischer E, Doniger R, Elson D, Chargaff E (1950). "The separation and estimation of ribonucleotides in minute quantities". *J Biol Chem* **186** (1): 37–50. PMID 14778802.
- [12] Chargaff E (1950). "Chemical specificity of nucleic acids and mechanism of their enzymatic degradation". *Experientia* **6** (6): 201–209.
- [13] Hosemann R., Bagchi R.N., *Direct analysis of diffraction by matter*, North-Holland Pubs., Amsterdam – New York, 1962.
- [14] Baianu, I.C. (1978). "X-ray scattering by partially disordered membrane systems.". *Acta Cryst.*, **A34** (5): 751–753. doi:10.1107/S0567739478001540.
- [15] http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=Abstract&list_uids=12379938
- [16] Mao, Chengde; Sun, Weiqiong & Seeman, Nadrian C. (16 June 1999). "Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy". *Journal of the American Chemical Society* **121** (23): 5437–5443. doi:10.1021/ja9900398. ISSN 0002-7863.
- [17] <http://ndbserver.rutgers.edu/atlas/xray/structures/U/ud0017/ud0017.html>
- [18] <http://www.phy.cam.ac.uk/research/bss/molbiophysics.php>
- [19] <http://planetphysics.org/encyclopedia/TheoreticalBiophysics.html>
- [20] (<http://www.jonathanpmiller.com/Karplus.html>)- obtaining dihedral angles from ³J coupling constants
- [21] (http://www.spectroscopynow.com/FCKeditor/UserFiles/File/specNOW/HTML files/General_Karplus_Calculator.htm) Another Javascript-like NMR coupling constant to dihedral
- [22] <http://ndbserver.rutgers.edu/atlas/nmr/index.html>
- [23] <http://ndbserver.rutgers.edu/ftp/NDB/coordinates/na-nmr-mmCIF/>
- [24] <http://ndbserver.rutgers.edu/ftp/NDB/nmr-restraints/>
- [25] Lee, S. C. et al., (2001). One Micrometer Resolution NMR Microscopy. *J. Magn. Res.*, **150**: 207-213.
- [26] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCs Press., Champaign, IL.
- [27] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy. 2004.I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004) (<http://arxiv.org/abs/q-bio/0407006>)
- [28] Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
- [29] [http://www.imaging.net/chemical-imaging/Chemical imaging](http://www.imaging.net/chemical-imaging/Chemical%20imaging)
- [30] http://www.malvern.com/LabEng/products/sdi/bibliography/sdi_bibliography.htm E. N. Lewis, E. Lee and L. H. Kidder, Combining Imaging and Spectroscopy: Solving Problems with Near-Infrared Chemical Imaging. *Microscopy Today*, Volume 12, No. 6, 11/2004.
- [31] D.S. Mantus and G. H. Morrison. 1991. Chemical imaging in biology and medicine using ion microscopy., *Microchimica Acta*, **104**, (1-6) January 1991, doi: 10.1007/BF01245536
- [32] J. Dubois, G. Sando, E. N. Lewis, Near-Infrared Chemical Imaging, A Valuable Tool for the Pharmaceutical Industry, G.I.T. Laboratory Journal Europe, No.1-2, 2007.
- [33] Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology.(June 2004).,I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin q-bio/0406047 (<http://arxiv.org/abs/q-bio/0406047>)
- [34] Chemical Imaging Without Dyeing (<http://witec.de/en/download/Raman/ImagingMicroscopy04.pdf>)
- [35] C.L. Evans and X.S. Xie.2008. Coherent Anti-Stokes Raman Scattering Microscopy: Chemical Imaging for Biology and Medicine., doi:10.1146/annurev.anchem.1.031207.112754 *Annual Review of Analytical Chemistry*, **1**: 883-909.
- [36] Eigen, M., Rigler, R. Sorting single molecules: application to diagnostics and evolutionary biotechnology,(1994) *Proc. Natl. Acad. Sci. USA*, 91,5740-5747.

- [37] Rigler, M. Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology, (1995) J. Biotechnol., 41,177-186.
- [38] Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.
- [39] Oehlenschläger F., Schwille P. and Eigen M. (1996). Detection of HIV-1 RNA by nucleic acid sequence-based amplification combined with fluorescence correlation spectroscopy, *Proc. Natl. Acad. Sci. USA* **93**:1281.
- [40] Bagatolli, L.A., and Gratton, E. (2000). Two-photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures. *Biophys J.*, 78:290-305.
- [41] Schwille, P., Haupts, U., Maiti, S., and Webb, W. (1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophysical Journal*, 77(10):2251-2265.
- [42] FRET description (<http://dwb.unl.edu/Teacher/NSF/C08/C08Links/pps99.cryst.bbk.ac.uk/projects/gmocz/fret.htm>)
- [43] doi:10.1016/S0959-440X(00)00190-1 ([http://dx.doi.org/10.1016/S0959-440X\(00\)00190-1](http://dx.doi.org/10.1016/S0959-440X(00)00190-1)) Recent advances in FRET: distance determination in protein–DNA complexes. *Current Opinion in Structural Biology* **2001**, 11(2), 201-207
- [44] <http://www.fretimaging.org/mcnamaraintro.html> FRET imaging introduction
- [45] <http://peds.oxfordjournals.org/cgi/content/abstract/1/4/295>
- [46] <http://www.sciencemag.org/cgi/content/abstract/sci;302/5649/1380>
- [47] <http://humphry.chem.wesleyan.edu:8080/MDDNA/>
- [48] <http://mmb.pcb.ub.es/DNAlive>
- [49] <http://diprodb.fli-leibniz.de>
- [50] <http://planetphysics.org/encyclopedia/BesselFunctionsApplicationsToDiffractionByHelicalStructures.html>
- [51] <http://planetphysics.org/?op=getobj&from=objects&name=BesselFunctionsAndTheirApplicationsToDiffractionByHelicalStructures>
- [52] <http://www.ntmdt.ru/SPM-Techniques/Principles/>
- [53] <http://ndbserver.rutgers.edu/atlas/xray/index.html>
- [54] <http://ndbserver.rutgers.edu/ftp/NDB/coordinates/na-biol/>
- [55] <http://ndbserver.rutgers.edu/ftp/NDB/structure-factors/>
- [56] <http://www.isis.rl.ac.uk/>
- [57] <http://www.cbs.dtu.dk/services/GenomeAtlas/>
- [58] <http://tubic.tju.edu.cn/zcurve/>
- [59] <http://ndbserver.rutgers.edu/ftp/NDB/models/>
- [60] http://www.parkafm.com/New_html/resources/01general.php
- [61] <http://www.rhk-tech.com/results/showcase.php>
-

DNA structure

The structure of DNA shows a variety of forms, both double-stranded and single-stranded. The mechanical properties of DNA, which are directly related to its structure, are a significant problem for cells. Every process which binds or reads DNA is able to use or modify the mechanical properties of DNA for purposes of recognition, packaging and modification. The extreme length (a chromosome may contain a 10 cm long DNA strand), relative rigidity and helical structure of DNA has led to the evolution of histones and of enzymes such as topoisomerases and helicases to manage a cell's DNA. The properties of DNA are closely related to its molecular structure and sequence, particularly the weakness of the hydrogen bonds and electronic interactions that hold strands of DNA together compared to the strength of the bonds within each strand.

Experimental techniques which can directly measure the mechanical properties of DNA are relatively new, and high-resolution visualization in solution is often difficult. Nevertheless, scientists have uncovered large amount of data on the mechanical properties of this polymer, and the implications of DNA's mechanical properties on cellular processes is a topic of active current research.

The DNA found in many cells can be macroscopic in length - a few centimetres long for each human chromosome. Consequently, cells must compact or "package" DNA to carry it within them. In eukaryotes this is carried by spool-like proteins known as histones, around which DNA winds. It is the further compaction of this DNA-protein complex which produces the well known mitotic eukaryotic chromosomes.

Structure determination

DNA structures can be determined using either nuclear magnetic resonance spectroscopy or X-ray crystallography. The first published reports of A-DNA X-ray diffraction patterns-- and also B-DNA—employed analyses based on Patterson transforms that provided only a limited amount of structural information for oriented fibers of DNA isolated from calf thymus.^{[1] [2]} An alternate analysis was then proposed by Wilkins et al. in 1953 for B-DNA X-ray diffraction/scattering patterns of hydrated, bacterial oriented DNA fibers and trout sperm heads in terms of squares of Bessel functions.^[3] Although the 'B-DNA form' is most common under the conditions found in cells,^[4] it is not a well-defined conformation but a family or fuzzy set of DNA-conformations that occur at the high hydration levels present in a wide variety of living cells.^[5] Their corresponding X-ray diffraction & scattering patterns are characteristic of molecular paracrystals with a significant degree of disorder (>20%)^{[6] [7]}, and concomitantly the structure is not tractable using only the standard analysis.

On the other hand, the standard analysis, involving only Fourier transforms of Bessel functions^[8] and DNA molecular models, is still routinely employed for the analysis of A-DNA and Z-DNA X-ray diffraction patterns.^[9]

Base pair geometry

The geometry of a base, or base pair step can be characterized by 6 coordinates: Shift, Slide, Rise, Tilt, Roll, and Twist. These values precisely define the location and orientation in space of every base or base pair in a DNA molecule relative to its predecessor along the axis of the helix. Together, they characterize the helical structure of the molecule. In regions of DNA where the "normal" structure is disrupted the change in these values can be used to describe such disruption.

For each base pair, considered relative to its predecessor^{[10] [11] [12]}:

- Shear
 - Stretch
 - Stagger
 - Buckle
-

- Propeller twist: *Rotation of one base with respect to the other in the same base pair.*
- Opening
- Shift: *displacement along an axis in the base-pair plane perpendicular to the first, directed from the minor to the major groove.*
- Slide: *displacement along an axis in the plane of the base pair directed from one strand to the other.*
- Rise: *displacement along the helix axis.*
- Tilt: *rotation around this axis.*
- Roll: *rotation around this axis.*
- Twist: *rotation around the helix axis.*
- x-displacement
- y-displacement
- inclination
- tip
- pitch: *the number of base pairs per complete turn of the helix*

Rise and twist determine the handedness and pitch of the helix. The other coordinates, by contrast, can be zero. Slide and shift are typically small in B-DNA, but are substantial in A- and Z-DNA. Roll and tilt make successive base pairs less parallel, and are typically small. A diagram ^[13] of these coordinates can be found in 3DNA ^[14] website.

Note that "tilt" has often been used differently in the scientific literature, referring to the deviation of the first, inter-strand base-pair axis from perpendicularity to the helix axis. This corresponds to slide between a succession of base pairs, and in helix-based coordinates is properly termed "inclination".

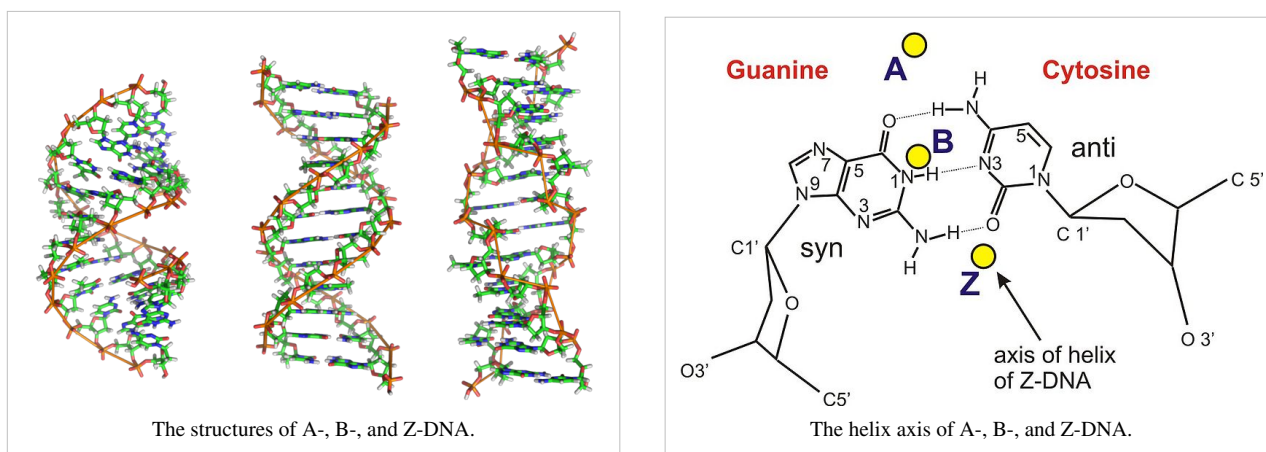
DNA helix geometries

Three DNA conformations are believed to be found in nature, A-DNA, B-DNA, and Z-DNA. The "B" form described by James D. Watson and Francis Crick is believed to predominate in cells ^[15]. It is 23.7 Å wide and extends 34 Å per 10 bp of sequence. The double helix makes one complete turn about its axis every 10.4-10.5 base pairs in solution. This frequency of twist (known as the helical *pitch*) depends largely on stacking forces that each base exerts on its neighbours in the chain.

Other conformations are possible; A-DNA, B-DNA, C-DNA, D-DNA ^[16], E-DNA ^[17], L-DNA (enantiomeric form of D-DNA) ^[16], P-DNA ^[18], S-DNA, Z-DNA, etc. have been described so far. ^[19] In fact, only the letters F, Q, U, V, and Y are now available to describe any new DNA structure that may appear in the future. ^[20] ^[21] However, most of these forms have been created synthetically and have not been observed in naturally occurring biological systems. Also note the triple-stranded DNA possibility.

A- and Z-DNA

A-DNA and Z-DNA differ significantly in their geometry and dimensions to B-DNA, although still form helical structures. The A form appears likely to occur only in dehydrated samples of DNA, such as those used in crystallographic experiments, and possibly in hybrid pairings of DNA and RNA strands. Segments of DNA that cells have methylated for regulatory purposes may adopt the Z geometry, in which the strands turn about the helical axis the opposite way to A-DNA and B-DNA. There is also evidence of protein-DNA complexes forming Z-DNA structures.



Structural features of the three major forms of DNA

Geometry attribute	A-DNA	B-DNA	Z-DNA
Helix sense	right-handed	right-handed	left-handed
Repeating unit	1 bp	1 bp	2 bp
Rotation/bp	33.6°	35.9°	60°/2bp
Mean bp/turn	10.7	10.0	12
Inclination of bp to axis	+19°	-1.2°	-9°
Rise/bp along axis	2.3 Å	3.32 Å	3.8 Å
Pitch/turn of helix	24.6 Å	33.2 Å	45.6 Å
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	anti	anti	C: anti, G: syn
Sugar pucker	C3'-endo	C2'-endo	C: C2'-endo, G: C2'-exo
Diameter	25.5 Å	23.7 Å	18.4 Å

Supercoiled DNA

The B form of the DNA helix twists 360° per 10.4-10.5 bp in the absence of torsional strain. But many molecular biological processes can induce torsional strain. A DNA segment with excess or insufficient helical twisting is referred to, respectively, as positively or negatively "supercoiled". DNA *in vivo* is typically negatively supercoiled, which facilitates the unwinding (melting) of the double-helix required for RNA transcription.

Non-helical forms

Other non-double helical forms of DNA have been described, for example side-by-side (SBS) and triple helical configurations. Single stranded DNA may exist *in statu nascendi* or as thermally induced despiralized DNA.

DNA bending

DNA is a relatively rigid polymer, typically modelled as a worm-like chain. It has three significant degrees of freedom; bending, twisting and compression, each of which cause particular limitations on what is possible with DNA within a cell. Twisting/torsional stiffness is important for the circularisation of DNA and the orientation of DNA bound proteins relative to each other and bending/axial stiffness is important for DNA wrapping and circularisation and protein interactions. Compression/extension is relatively unimportant in the absence of high tension.

Persistence length/Axial stiffness

Example sequences and their persistence lengths (B DNA)

Sequence	Persistence Length /base pairs
Random	154±10
(CA) _{repeat}	133±10
(CAG) _{repeat}	124±10
(TATA) _{repeat}	137±10

DNA in solution does not take a rigid structure but is continually changing conformation due to thermal vibration and collisions with water molecules, which makes classical measures of rigidity impossible. Hence, the bending stiffness of DNA is measured by the persistence length, defined as:

"The length of DNA over which the time-averaged orientation of the polymer becomes uncorrelated by a factor of e ."

This value may be directly measured using an atomic force microscope to directly image DNA molecules of various lengths. In aqueous solution the average persistence length is 46-50 nm or 140-150 base pairs (the diameter of DNA is 2 nm), although can vary significantly. This makes DNA a moderately stiff molecule.

The persistence length of a section of DNA is somewhat dependent on its sequence, and this can cause significant variation. The variation is largely due to base stacking energies and the residues which extend into the minor and major grooves.

Models for DNA bending

Stacking stability of base steps (B DNA)

Step	Stacking ΔG /kcal mol ⁻¹
T A	-0.19
T G or C A	-0.55
C G	-0.91
A G or C T	-1.06
A A or T T	-1.11
A T	-1.34

G A or T C	-1.43
C C or G G	-1.44
A C or G T	-1.81
G C	-2.17

The entropic flexibility of DNA is remarkably consistent with standard polymer physics models such as the *Kratky-Porod* worm-like chain model. Consistent with the worm-like chain model is the observation that bending DNA is also described by Hooke's law at very small (sub-piconewton) forces. However for DNA segments less than the persistence length, the bending force is approximately constant and behaviour deviates from the worm-like chain predictions.

This effect results in unusual ease in circularising small DNA molecules and a higher probability of finding highly bent sections of DNA.

Bending preference

DNA molecules often have a preferred direction to bend, ie. anisotropic bending. This is, again, due to the properties of the bases which make up the DNA sequence - a random sequence will have no preferred bend direction, i.e. isotropic bending.

Preferred DNA bend direction is determined by the stability of stacking each base on top of the next. If unstable base stacking steps are always found on one side of the DNA helix then the DNA will preferentially bend away from that direction. As bend angle increases then steric hindrances and ability to roll the residues relative to each other also play a role, especially in the minor groove. **A** and **T** residues will be preferentially be found in the minor grooves on the inside of bends. This effect is particularly seen in DNA-protein binding where tight DNA bending is induced, such as in nucleosome particles. See base step distortions above.

DNA molecules with exceptional bending preference can become intrinsically bent. This was first observed in trypanosomatid kinetoplast DNA. Typical sequences which cause this contain stretches of 4-6 **T** and **A** residues separated by **G** and **C** rich sections which keep the A and T residues in phase with the minor groove on one side of the molecule. For example:

The diagram shows a DNA sequence with a highlighted region. The sequence is as follows:

Top strand: G A T T C C C A A A A T G T C A A A A A A T A G G C A A A A A A T G C

Bottom strand: C A A A A A T C C C A A A C

The highlighted region is the sequence: A A A A T G T C A A A A A A T A G G C A A A A A A T G C (top strand) and T C C C A A A C (bottom strand).

A vertical line marks a mutation site at the 14th position of the top strand, where the sequence changes from A A A A A A T A G G C to A A A A A A T G C.

The intrinsically bent structure is induced by the 'propeller twist' of base pairs relative to each other allowing unusual bifurcated Hydrogen-bonds between base steps. At higher temperatures this structure, and so the intrinsic bend, is lost.

All DNA which bends anisotropically has, on average, a longer persistence length and greater axial stiffness. This increased rigidity is required to prevent random bending which would make the molecule act isotropically.

DNA circularization

DNA circularization depends on both the axial (bending) stiffness and torsional (rotational) stiffness of the molecule. For a DNA molecule to successfully circularize it must be long enough to easily bend into the full circle and must have the correct number of bases so the ends are in the correct rotation to allow bonding to occur. The optimum length for circularization of DNA is around 400 base pairs (136 nm), with an integral number of turns of the DNA helix, i.e. multiples of 10.4 base pairs. Having a non integral number of turns presents a significant energy barrier for circularization, for example a $10.4 \times 30 = 312$ base pair molecule will circularize hundreds of times faster than $10.4 \times$

30.5 $\approx$ 317 base pair molecule.

DNA stretching

Longer stretches of DNA are entropically elastic under tension. When DNA is in solution, it undergoes continuous structural variations due to the energy available in the solvent. This is due to the thermal vibration of the molecule combined with continual collisions with water molecules. For entropic reasons, more compact relaxed states are thermally accessible than stretched out states, and so DNA molecules are almost universally found in a tangled relaxed layouts. For this reason, a single molecule of DNA will stretch under a force, straightening it out. Using optical tweezers, the entropic stretching behavior of DNA has been studied and analyzed from a polymer physics perspective, and it has been found that DNA behaves largely like the *Kratky-Porod* worm-like chain model under physiologically accessible energy scales.

Under sufficient tension and positive torque, DNA is thought to undergo a phase transition with the bases splaying outwards and the phosphates moving to the middle. This proposed structure for overstretched DNA has been called "P-form DNA," in honor of Linus Pauling who originally presented it as a possible structure of DNA^[18]

The mechanical properties DNA under compression have not been characterized due to experimental difficulties in preventing the polymer from bending under the compressive force.

DNA melting

Melting stability of base steps (B DNA)

Step	Melting ΔG /Kcal mol ⁻¹
T A	-0.12
T G or C A	-0.78
C G	-1.44
A G or C T	-1.29
A A or T T	-1.04
A T	-1.27
G A or T C	-1.66
C C or G G	-1.97
A C or G T	-2.04
G C	-2.70

DNA melting is the process by which the interactions between the strands of the double helix are broken, separating the two strands of DNA. These bonds are weak, easily separated by gentle heating, enzymes, or physical force. DNA melting preferentially occurs at certain points in the DNA.^[22] **T** and **A** rich sequences are more easily melted than **C** and **G** rich regions. Particular base steps are also susceptible to DNA melting, particularly **T A** and **T G** base steps.^[23] These mechanical features are reflected by the use of sequences such as **TATAA** at the start of many genes to assist RNA polymerase in melting the DNA for transcription.

Strand separation by gentle heating, as used in PCR, is simple providing the molecules have fewer than about 10,000 base pairs (10 kilobase pairs, or 10 kbp). The intertwining of the DNA strands makes long segments difficult to separate. The cell avoids this problem by allowing its DNA-melting enzymes (helicases) to work concurrently with

topoisomerases, which can chemically cleave the phosphate backbone of one of the strands so that it can swivel around the other. Helicases unwind the strands to facilitate the advance of sequence-reading enzymes such as DNA polymerase.

DNA topology

Within the cell most DNA is topologically restricted. DNA is typically found in closed loops (such as plasmids in prokaryotes) which are topologically closed, or as very long molecules whose diffusion coefficients produce effectively topologically closed domains. Linear sections of DNA are also commonly bound to proteins or physical structures (such as membranes) to form closed topological loops.

Francis Crick was one of the first to propose the importance of linking numbers when considering DNA supercoils. In a paper published in 1976, Crick outlined the problem as follows:

In considering supercoils formed by closed double-stranded molecules of DNA certain mathematical concepts, such as the linking number and the twist, are needed. The meaning of these for a closed ribbon is explained and also that of the writhing number of a closed curve. Some simple examples are given, some of which may be relevant to the structure of chromatin.^[24]

Analysis of DNA topology uses three values:

L = linking number - the number of times one DNA strand wraps around the other. It is an integer for a closed loop and constant for a closed topological domain.

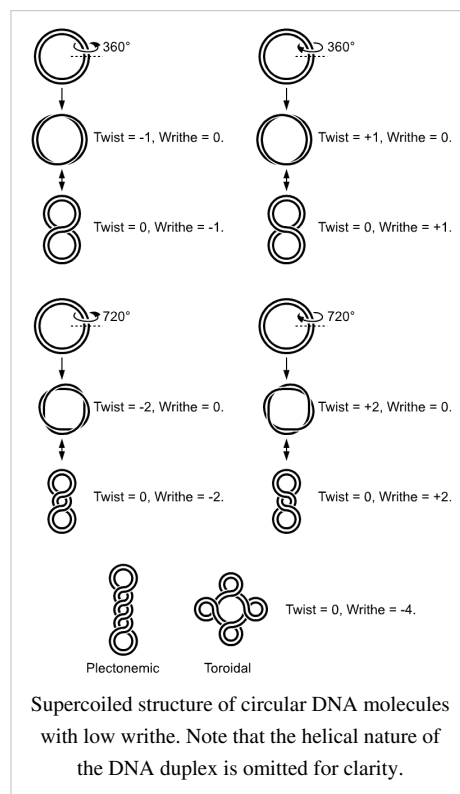
T = twist - total number of turns in the double stranded DNA helix. This will normally try to be equal to the number turns a DNA molecule will make while free in solution, ie. number of bases/10.4.

W = writhe - number of turns of the double stranded DNA helix around the superhelical axis

$$L = T + W \text{ and } \Delta L = \Delta T + \Delta W$$

Any change of T in a closed topological domain must be balanced by a change in W , and vice versa. This results in higher order structure of DNA. A circular DNA molecule with a writhe of 0 will be circular. If the twist of this molecule is subsequently increased or decreased by supercoiling then the writhe will be appropriately altered, making the molecule undergo plectonemic or toroidal superhelical coiling.

When the ends of a piece of double stranded helical DNA are joined so that it forms a circle the strands are topologically knotted. This means the single strands cannot be separated any process that does not involve breaking a strand (such as heating). The task of un-knotting topologically linked strands of DNA falls to enzymes known as topoisomerases. These enzymes are dedicated to un-knotting circular DNA by cleaving one or both strands so that another double or single stranded segment can pass through. This un-knotting is required for the replication of circular DNA and various types of recombination in linear DNA which have similar topological constraints.



The linking number paradox

For many years, the origin of residual supercoiling in eukaryotic genomes remained unclear. This topological puzzle was referred to by some as the "linking number paradox".^[25] However, when experimentally determined structures of the nucleosome displayed an overtwisted left-handed wrap of DNA around the histone octamer^{[26] [27]}, this "paradox" was solved.

See also

- DNA nanotechnology
- Molecular models of DNA

External links

- MDDNA: Structural Bioinformatics of DNA^[47]
- Abalone^[28] — Commercial software for DNA modeling
- DNALive: a web interface to compute DNA physical properties^[48]. Also allows cross-linking of the results with the UCSC Genome browser and DNA dynamics.
- DiProDB: Dinucleotide Property Database^[49]. The database is designed to collect and analyse thermodynamic, structural and other dinucleotide properties.

References

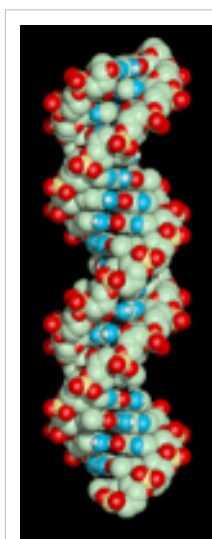
- [1] Franklin, R.E. and Gosling, R.G. received 6 March 1953. *Acta Cryst.* (1953). **6**, 673: The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content.; also *Acta Cryst.* **6**, 678: The Structure of Sodium Thymonucleate Fibres II. The Cylindrically Symmetrical Patterson Function.
- [2] Franklin, Rosalind (1953). "Molecular Configuration in Sodium Thymonucleate. Franklin R. and Gosling R.G" (<http://www.nature.com/nature/dna50/franklingosling.pdf>) (PDF). *Nature* **171**: 740–741. doi:10.1038/171740a0. PMID 13054694. .
- [3] Wilkins M.H.F., A.R. Stokes A.R. & Wilson, H.R. (1953). "Molecular Structure of Deoxypentose Nucleic Acids" (<http://www.nature.com/nature/dna50/wilkins.pdf>) (PDF). *Nature* **171**: 738–740. doi:10.1038/171738a0. PMID 13054693. .
- [4] Leslie AG, Arnott S, Chandrasekaran R, Ratliff RL (1980). "Polymorphism of DNA double helices". *J. Mol. Biol.* **143** (1): 49–72. doi:10.1016/0022-2836(80)90124-2. PMID 7441761.
- [5] Baianu, I.C. (1980). "Structural Order and Partial Disorder in Biological systems". *Bull. Math. Biol.* **42** (4): 464–468. doi:10.1016/0022-2836(80)90124-2.
- [6] Hosemann R., Bagchi R.N., *Direct analysis of diffraction by matter*, North-Holland Publs., Amsterdam – New York, 1962
- [7] Baianu I.C., X-ray scattering by partially disordered membrane systems, *Acta Cryst. A*, **34** (1978), 751–753.
- [8] Bessel functions and diffraction by helical structures (<http://planetphysics.org/encyclopedia/BesselFunctionsAndTheirApplicationsToDiffractionByHelicalStructures.html>)
- [9] X-Ray Diffraction Patterns of Double-Helical Deoxyribonucleic Acid (DNA) Crystals (<http://planetphysics.org/encyclopedia/BesselFunctionsApplicationsToDiffractionByHelicalStructures.html>)
- [10] Dickerson RE (1989). "Definitions and nomenclature of nucleic acid structure components". *Nucleic Acids Res* **17** (5): 1797–1803. doi:10.1093/nar/17.5.1797. PMID 2928107.
- [11] Lu XJ, Olson WK (1999). "Resolving the discrepancies among nucleic acid conformational analyses". *J Mol Biol* **285** (4): 1563–1575. doi:10.1006/jmbi.1998.2390. PMID 9917397.
- [12] Olson WK, Bansal M, Burley SK, Dickerson RE, Gerstein M, Harvey SC, Heinemann U, Lu XJ, Neidle S, Shakked Z, Sklenar H, Suzuki M, Tung CS, Westhof E, Wolberger C, Berman HM (2001). "A standard reference frame for the description of nucleic acid base-pair geometry". *J Mol Biol* **313** (1): 229–237. doi:10.1006/jmbi.2001.4987. PMID 11601858.
- [13] http://rutchem.rutgers.edu/~xiangjun/3DNA/images/bp_step_hel.gif
- [14] <http://rutchem.rutgers.edu/~xiangjun/3DNA/examples.html>
- [15] Richmond, et al. (2003). "The structure of DNA in the nucleosome core". *Nature* **423**: 145–150. doi:10.1038/nature01595. PMID 12736678.
- [16] Hayashi G, Hagihara M, Nakatani K (2005). "Application of L-DNA as a molecular tag". *Nucleic Acids Symp Ser (Oxf)* **49**: 261–262. PMID 17150733.
- [17] Vargason JM, Eichman BF, Ho PS (2000). "The extended and eccentric E-DNA structure induced by cytosine methylation or bromination". *Nature Structural Biology* **7**: 758–761. doi:10.1038/78985.
- [18] Allemand, et al. (1998). "Stretched and overwound DNA forms a Pauling-like structure with exposed bases". *PNAS* **24**: 14152–14157. doi:10.1073/pnas.95.24.14152. PMID 9826669.

- [19] List of 55 fiber structures (http://rutchem.rutgers.edu/~xiangjun/3DNA/misc/fiber_model.txt)
- [20] Bansal M (2003). "DNA structure: Revisiting the Watson-Crick double helix". *Current Science* **85** (11): 1556–1563.
- [21] Ghosh A, Bansal M (2003). "A glossary of DNA structures from A to Z". *Acta Cryst* **D59**: 620–626. doi:10.1107/S0907444903003251.
- [22] Breslauer KJ, Frank R, Blöcker H, Marky LA (1986). "Predicting DNA duplex stability from the base sequence". *PNAS* **83** (11): 3746–3750. PMID 3459152.
- [23] Richard Owczarzy (2008-08-28). "DNA melting temperature - How to calculate it?" (<http://www.owczarzy.net/tm.htm>). *High-throughput DNA biophysics*. owczarzy.net. . Retrieved 2008-10-02.
- [24] Crick FH (1976). "Linking numbers and nucleosomes". *Proc Natl Acad Sci USA* **73** (8): 2639–43. doi:10.1073/pnas.73.8.2639. PMID 1066673.
- [25] Prunell A (1998). "A topological approach to nucleosome structure and dynamics: the linking number paradox and other issues". *Biophys J* **74** (5): 2531–2544. PMID 9591679.
- [26] Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997). "Crystal structure of the nucleosome core particle at 2.8 Å resolution". *Nature* **389** (6648): 251–260. doi:10.1038/38444. PMID 9305837.
- [27] Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ (2002). "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution". *Journal of Molecular Biology* **319** (5): 1097–1113. doi:10.1016/S0022-2836(02)00386-8. PMID 12079350.
- [28] <http://www.biomolecular-modeling.com/Abalone/index.html>

DNA Dynamics

Molecular models of DNA structures are representations of the molecular geometry and topology of Deoxyribonucleic acid (DNA) molecules using one of several means, with the aim of simplifying and presenting the essential, physical and chemical, properties of DNA molecular structures either *in vivo* or *in vitro*. These representations include closely packed spheres (CPK models) made of plastic, metal wires for 'skeletal models', graphic computations and animations by computers, artistic rendering. Computer molecular models also allow animations and molecular dynamics simulations that are very important for understanding how DNA functions *in vivo*.

The more advanced, computer-based molecular models of DNA involve molecular dynamics simulations as well as quantum mechanical computations of vibro-rotations, delocalized molecular orbitals (MOs), electric dipole moments, hydrogen-bonding, and so on. **DNA molecular dynamics modeling** involves simulations of DNA molecular geometry and topology changes with time as a result of both intra- and inter- molecular interactions of DNA. Whereas molecular models of Deoxyribonucleic acid (DNA) molecules such as closely packed spheres (CPK models) made of plastic or metal wires for 'skeletal models' are useful representations of static DNA structures, their usefulness is very limited for representing complex DNA dynamics. Computer molecular modeling allows both animations and molecular dynamics simulations that are very important for understanding how DNA functions *in vivo*.



Spinning DNA
generic model.

History

Double Helix Discovery



William Astbury

Oswald Avery

Francis Crick

Erwin Chargaff

Max Delbrück

Jerry Donohue

Rosalind Franklin

Raymond Gosling

Phoebus Levene

Linus Pauling

Sir John Randall

Erwin Schrödinger

Alex Stokes

James Watson

Maurice Wilkins

Herbert Wilson

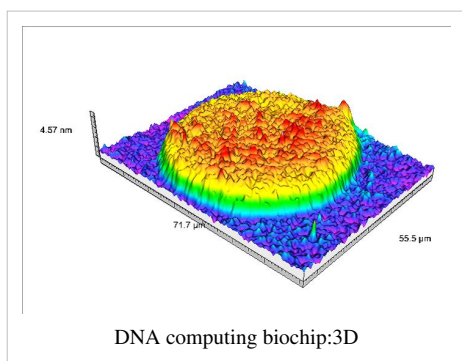
From the very early stages of structural studies of DNA by X-ray diffraction and biochemical means, molecular models such as the Watson-Crick double-helix model were successfully employed to solve the 'puzzle' of DNA structure, and also find how the latter relates to its key functions in living cells. The first high quality X-ray diffraction patterns of A-DNA were reported by Rosalind Franklin and Raymond Gosling in 1953^[1]. The first calculations of the Fourier transform of an atomic helix were reported one year earlier by Cochran, Crick and Vand^[2], and were followed in 1953 by the computation of the Fourier transform of a coiled-coil by Crick^[3].

Structural information is generated from X-ray diffraction studies of oriented DNA fibers with the help of molecular models of DNA that are combined with crystallographic and mathematical analysis of the X-ray patterns.

The first reports of a double-helix molecular model of B-DNA structure were made by Watson and Crick in 1953^[4]^[5]. Last-but-not-least, Maurice F. Wilkins, A. Stokes and H.R. Wilson, reported the first X-ray patterns of *in vivo* B-DNA in partially oriented salmon sperm heads^[6]. The development of the first correct double-helix molecular model of DNA by Crick and Watson may not have been possible without the biochemical evidence for the nucleotide base-pairing ([A---T]; [C---G]), or Chargaff's rules^[7] [8] [9] [10] [11] [12]. Although such initial studies of DNA structures with the help of molecular models were essentially static, their consequences for explaining the *in vivo* functions of DNA were significant in the areas of protein biosynthesis and the quasi-universality of the genetic code. Epigenetic transformation studies of DNA *in vivo* were however much slower to develop in spite of their importance for embryology, morphogenesis and cancer research. Such chemical dynamics and biochemical reactions of DNA are much more complex than the molecular dynamics of DNA physical interactions with water, ions and proteins/enzymes in living cells.

Importance

An old standing dynamic problem is how DNA "self-replication" takes place in living cells that should involve transient uncoiling of supercoiled DNA fibers. Although DNA consists of relatively rigid, very large elongated biopolymer molecules called "fibers" or chains (that are made of repeating nucleotide units of four basic types, attached to deoxyribose and phosphate groups), its molecular structure *in vivo* undergoes dynamic configuration changes that involve dynamically attached water molecules and ions. Supercoiling, packing with histones in chromosome structures, and other such supramolecular aspects also involve *in vivo* DNA topology which is even more complex than DNA molecular geometry, thus turning molecular modeling of DNA into an especially challenging problem for both molecular biologists and biotechnologists. Like other large molecules and biopolymers, DNA often exists in multiple stable geometries (that is, it exhibits conformational isomerism) and configurational, quantum states which are close to each other in energy on the potential energy surface of the DNA molecule.



Such varying molecular geometries can also be computed, at least in principle, by employing *ab initio* quantum chemistry methods that can attain high accuracy for small molecules, although claims that acceptable accuracy can be also achieved for polynucleotides, as well as DNA conformations, were recently made on the basis of VCD spectral data. Such quantum geometries define an important class of *ab initio* molecular models of DNA whose exploration has barely started especially in connection with results obtained by VCD in solutions. More detailed comparisons with such *ab initio* quantum computations are in principle obtainable through 2D-FT NMR spectroscopy and relaxation studies of polynucleotide solutions or specifically labeled DNA, as for example with deuterium labels.

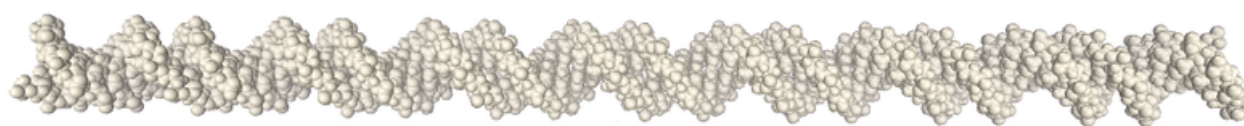
In an interesting twist of roles, the DNA molecule itself was proposed to be utilized for quantum computing. Both DNA nanostructures as well as DNA 'computing' biochips have been built (see biochip image at left).

Examples of DNA molecular models

Animated molecular models allow one to visually explore the three-dimensional (3D) structure of DNA. One visualization of DNA model is a space-filling, or CPK, model. Another is a wire, or skeletal type.

The hydrogen bonding dynamics and proton exchange is very different by many orders of magnitude between the two systems of fully hydrated DNA and water molecules in ice. Thus, the DNA dynamics is complex, involving nanosecond and several tens of picosecond time scales, whereas that of liquid ice is on the picosecond time scale, and that of proton exchange in ice is on the millisecond time scale; the proton exchange rates in DNA and attached proteins may vary from picosecond to nanosecond, minutes or years, depending on the exact locations of the exchanged protons in the large biopolymers.

A simple harmonic oscillator 'vibration' is only an oversimplified dynamic representation of the longitudinal vibrations of the DNA intertwined helices which were found to be anharmonic rather than harmonic as often assumed in quantum dynamic simulations of DNA.



DNA Spacefilling molecular model

Paracrystalline lattice models of B-DNA structures

A paracrystalline lattice, or paracrystal, is a molecular or atomic lattice with significant amounts (e.g., larger than a few percent) of partial disordering of molecular arrangements. Limiting cases of the paracrystal model are nanostructures, such as glasses, liquids, etc., that may possess only local ordering and no global order. Liquid crystals also have paracrystalline rather than crystalline structures.

Highly hydrated B-DNA occurs naturally in living cells in such a paracrystalline state, which is a dynamic one in spite of the relatively rigid DNA double-helix stabilized by parallel hydrogen bonds between the nucleotide base-pairs in the two complementary, helical DNA chains (see figures). For simplicity most DNA molecular models omit both water and ions dynamically bound to B-DNA, and are thus less useful for understanding the dynamic behaviors of B-DNA *in vivo*. The physical and mathematical analysis of X-ray^[13] ^[14] and spectroscopic data for paracrystalline B-DNA is therefore much more complicated than that of crystalline, A-DNA X-ray diffraction patterns. The paracrystal model is also important for DNA technological applications such as DNA nanotechnology. Novel techniques that combine X-ray diffraction of DNA with X-ray microscopy in hydrated living cells are now also being developed (see, for example, "Application of X-ray microscopy in the analysis of living hydrated cells"^[15]).

Genomic and biotechnology applications of DNA molecular modeling

There are various uses of DNA molecular modeling in Genomics and Biotechnology research applications, from DNA repair to PCR and DNA nanostructures. Two-dimensional DNA junction arrays have been visualized by Atomic force microscopy.^[15]

Quadruplex DNA^[17] may be involved in certain cancers^[16] ^[17].

See also

- DNA structure
- X-ray scattering
- Neutron scattering
- Crystallography
- Crystal lattices
- 2D-FT NMRI and Spectroscopy
- List of nucleic acid simulation software
- Sirius visualization software
- X-ray microscopy
- QMC@Home
- Sir Lawrence Bragg, FRS

Spectroscopy

- Vibrational circular dichroism (VCD)
- FT-NMR^[18] ^[19]
 - NMR Atlas--database^[22]
 - mmCIF downloadable coordinate files of nucleic acids in solution from 2D-FT NMR data^[23]
 - NMR constraints files for NAs in PDB format^[24]
- NMR microscopy^[20]
- Microwave spectroscopy
- FT-IR
- FT-NIR^[21] ^[22] ^[23]

- Spectral, Hyperspectral, and Chemical imaging)^{[24] [25] [26] [21] [22] [27] [28]}.
- Raman spectroscopy/microscopy^[29] and CARS^[30].
- Fluorescence correlation spectroscopy^{[31] [32] [33] [22] [34] [35] [36] [21]}, Fluorescence cross-correlation spectroscopy and FRET^{[37] [38] [39]}.
- Confocal microscopy^[31]

Further reading

- *Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology*. (June 2004) I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin., q-bio/0406047.
- F. Bessel, *Untersuchung des Theils der planetarischen Störungen*, Berlin Abhandlungen (1824), article 14.
- Sir Lawrence Bragg, FRS. *The Crystalline State, A General survey*. London: G. Bells and Sons, Ltd., vols. 1 and 2., 1966., 2024 pages.
- Cantor, C. R. and Schimmel, P.R. *Biophysical Chemistry, Parts I and II.*, San Franscisco: W.H. Freeman and Co. 1980. 1,800 pages.
- Voet, D. and J.G. Voet. *Biochemistry*, 2nd Edn., New York, Toronto, Singapore: John Wiley & Sons, Inc., 1995, ISBN 0-471-58651-X., 1361 pages.
- Watson, G. N. *A Treatise on the Theory of Bessel Functions.*, (1995) Cambridge University Press. ISBN 0-521-48391-3.
- Watson, James D. *Molecular Biology of the Gene*. New York and Amsterdam: W.A. Benjamin, Inc. 1965., 494 pages.
- Wentworth, W.E. *Physical Chemistry. A short course.*, Malden (Mass.): Blackwell Science, Inc. 2000.
- Herbert R. Wilson, FRS. *Diffraction of X-rays by proteins, Nucleic Acids and Viruses.*, London: Edward Arnold (Publishers) Ltd. 1966.
- Kurt Wuthrich. *NMR of Proteins and Nucleic Acids.*, New York, Brisbane,Chicester, Toronto, Singapore: J. Wiley & Sons. 1986., 292 pages.
- Robinson, Bruche H.; Seeman, Nadrian C. (August 1987). "The Design of a Biochip: A Self-Assembling Molecular-Scale Memory Device". *Protein Engineering* **1** (4): 295–300. ISSN 0269-2139. Link ^[45]
- Rothmund, Paul W. K.; Ekani-Nkodo, Axel; Papadakis, Nick; Kumar, Ashish; Fygenson, Deborah Kuchnir & Winfree, Erik (22 December 2004). "Design and Characterization of Programmable DNA Nanotubes". *Journal of the American Chemical Society* **126** (50): 16344–16352. doi:10.1021/ja044319l. ISSN 0002-7863.
- Keren, K.; Kinneret Keren, Rotem S. Berman, Evgeny Buchstab, Uri Sivan, Erez Braun (November 2003). "DNA-Templated Carbon Nanotube Field-Effect Transistor" ^[46]. *Science* **302** (6549): 1380–1382. doi:10.1126/science.1091022. ISSN 1095-9203.
- Zheng, Jiwen; Constantinou, Pamela E.; Micheel, Christine; Alivisatos, A. Paul; Kiehl, Richard A. & Seeman Nadrian C. (2006). "2D Nanoparticle Arrays Show the Organizational Power of Robust DNA Motifs". *Nano Letters* **6**: 1502–1504. doi:10.1021/nl060994c. ISSN 1530-6984.
- Cohen, Justin D.; Sadowski, John P.; Dervan, Peter B. (2007). "Addressing Single Molecules on DNA Nanostructures". *Angewandte Chemie* **46** (42): 7956–7959. doi:10.1002/anie.200702767. ISSN 0570-0833.
- Constantinou, Pamela E.; Wang, Tong; Kopatsch, Jens; Israel, Lisa B.; Zhang, Xiaoping; Ding, Baoquan; Sherman, William B.; Wang, Xing; Zheng, Jianping; Sha, Ruojie & Seeman, Nadrian C. (2006). "Double cohesion in structural DNA nanotechnology". *Organic and Biomolecular Chemistry* **4**: 3414–3419. doi:10.1039/b605212f.
- Hallin PF, David Ussery D (2004). "CBS Genome Atlas Database: A dynamic storage for bioinformatic results and DNA sequence data". *Bioinformatics* **20**: 3682–3686.
- Zhang CT, Zhang R, Ou HY (2003). "The Z curve database: a graphic representation of genome sequences". *Bioinformatics* **19** (5): 593-599. doi:10.1093/bioinformatics/btg041

External links

- DNA the Double Helix Game ^[152] From the official Nobel Prize web site
- MDDNA: Structural Bioinformatics of DNA ^[47]
- Double Helix 1953–2003 ^[156] National Centre for Biotechnology Education
- DNALive: a web interface to compute DNA physical properties ^[48]. Also allows cross-linking of the results with the UCSC Genome browser and DNA dynamics.
- DiProDB: Dinucleotide Property Database ^[49]. The database is designed to collect and analyse thermodynamic, structural and other dinucleotide properties.
- Further details of mathematical and molecular analysis of DNA structure based on X-ray data ^[50]
- Bessel functions corresponding to Fourier transforms of atomic or molecular helices. ^[51]
- Application of X-ray microscopy in analysis of living hydrated cells ^[15]
- overview of STM/AFM/SNOM principles with educative videos ^[52]

Databases for DNA molecular models and sequences

X-ray diffraction

- NDB ID: UD0017 Database ^[17]
- X-ray Atlas -database ^[53]
- PDB files of coordinates for nucleic acid structures from X-ray diffraction by NA (incl. DNA) crystals ^[54]
- Structure factors downloadable files in CIF format ^[55]

Neutron scattering

- ISIS neutron source: ISIS pulsed neutron source: A world centre for science with neutrons & muons at Harwell, near Oxford, UK. ^[56]

X-ray microscopy

- Application of X-ray microscopy in the analysis of living hydrated cells ^[15]

Electron microscopy

- DNA under electron microscope ^[153]

Genomic and structural databases

- CBS Genome Atlas Database ^[57] — contains examples of base skews.
- The Z curve database of genomes — a 3-dimensional visualization and analysis tool of genomes ^[58].
- DNA and other nucleic acids' molecular models: Coordinate files of nucleic acids molecular structure models in PDB and CIF formats ^[59]

Atomic force microscopy

- How SPM Works ^[60]
 - SPM Image Gallery - AFM STM SEM MFM NSOM and more. ^[61]
-

References

- [1] Franklin, R.E. and Gosling, R.G. recd.6 March 1953. *Acta Cryst.* (1953). 6, 673 The Structure of Sodium Thymonucleate Fibres I. The Influence of Water Content *Acta Cryst.* (1953). and 6, 678 The Structure of Sodium Thymonucleate Fibres II. The Cylindrically Symmetrical Patterson Function.
- [2] Cochran, W., Crick, F.H.C. and Vand V. 1952. The Structure of Synthetic Polypeptides. 1. The Transform of Atoms on a Helix. *Acta Cryst.* 5(5):581-586.
- [3] Crick, F.H.C. 1953a. The Fourier Transform of a Coiled-Coil., *Acta Crystallographica* 6(8-9):685-689.
- [4] Watson, James D. and Francis H.C. Crick. A structure for Deoxyribose Nucleic Acid (<http://www.nature.com/nature/dna50/watsoncrick.pdf>) (PDF). *Nature* 171, 737–738, 25 April 1953.
- [5] Watson, J.D; Crick F.H.C. 1953b. The Structure of DNA., Cold Spring Harbor Symposia on Quantitative Biology 18:123-131.
- [6] Wilkins M.H.F., A.R. Stokes A.R. & Wilson, H.R. (1953). "Molecular Structure of Deoxypentose Nucleic Acids" (<http://www.nature.com/nature/dna50/wilkins.pdf>) (PDF). *Nature* **171** (4356): 738–740. doi:10.1038/171738a0. PMID 13054693. .
- [7] Elson D, Chargaff E (1952). "On the deoxyribonucleic acid content of sea urchin gametes". *Experientia* **8** (4): 143–145.
- [8] Chargaff E, Lipshitz R, Green C (1952). "Composition of the deoxypentose nucleic acids of four genera of sea-urchin". *J Biol Chem* **195** (1): 155–160. PMID 14938364.
- [9] Chargaff E, Lipshitz R, Green C, Hodes ME (1951). "The composition of the deoxyribonucleic acid of salmon sperm". *J Biol Chem* **192** (1): 223–230. PMID 14917668.
- [10] Chargaff E (1951). "Some recent studies on the composition and structure of nucleic acids". *J Cell Physiol Suppl* **38** (Suppl).
- [11] Magasanik B, Vischer E, Doniger R, Elson D, Chargaff E (1950). "The separation and estimation of ribonucleotides in minute quantities". *J Biol Chem* **186** (1): 37–50. PMID 14778802.
- [12] Chargaff E (1950). "Chemical specificity of nucleic acids and mechanism of their enzymatic degradation". *Experientia* **6** (6): 201–209.
- [13] Hosemann R., Bagchi R.N., *Direct analysis of diffraction by matter*, North-Holland Pubs., Amsterdam – New York, 1962.
- [14] Baianu, I.C. (1978). "X-ray scattering by partially disordered membrane systems.". *Acta Cryst.*, **A34** (5): 751–753. doi:10.1107/S0567739478001540.
- [15] Mao, Chengde; Sun, Weiqiong & Seeman, Nadrian C. (16 June 1999). "Designed Two-Dimensional DNA Holliday Junction Arrays Visualized by Atomic Force Microscopy". *Journal of the American Chemical Society* **121** (23): 5437–5443. doi:10.1021/ja9900398. ISSN 0002-7863.
- [16] <http://www.phy.cam.ac.uk/research/bss/molbiophysics.php>
- [17] <http://planetphysics.org/encyclopedia/TheoreticalBiophysics.html>
- [18] (<http://www.jonathanpmiller.com/Karplus.html>)- obtaining dihedral angles from ³J coupling constants
- [19] (http://www.spectroscopynow.com/FCkEditor/UserFiles/File/specNOW/HTML files/General_Karplus_Calculator.htm) Another Javascript-like NMR coupling constant to dihedral
- [20] Lee, S. C. et al., (2001). One Micrometer Resolution NMR Microscopy. *J. Magn. Res.*, **150**: 207-213.
- [21] Near Infrared Microspectroscopy, Fluorescence Microspectroscopy, Infrared Chemical Imaging and High Resolution Nuclear Magnetic Resonance Analysis of Soybean Seeds, Somatic Embryos and Single Cells., Baianu, I.C. et al. 2004., In *Oil Extraction and Analysis.*, D. Luthria, Editor pp.241-273, AOCS Press., Champaign, IL.
- [22] Single Cancer Cell Detection by Near Infrared Microspectroscopy, Infrared Chemical Imaging and Fluorescence Microspectroscopy. 2004.I. C. Baianu, D. Costescu, N. E. Hofmann, S. S. Korban and et al., q-bio/0407006 (July 2004) (<http://arxiv.org/abs/q-bio/0407006>)
- [23] Raghavachari, R., Editor. 2001. *Near-Infrared Applications in Biotechnology*, Marcel-Dekker, New York, NY.
- [24] [http://www.imaging.net/chemical-imaging/Chemical imaging](http://www.imaging.net/chemical-imaging/Chemical%20imaging)
- [25] http://www.malvern.com/LabEng/products/sdi/bibliography/sdi_bibliography.htm E. N. Lewis, E. Lee and L. H. Kidder, Combining Imaging and Spectroscopy: Solving Problems with Near-Infrared Chemical Imaging. *Microscopy Today*, Volume 12, No. 6, 11/2004.
- [26] D.S. Mantus and G. H. Morrison. 1991. Chemical imaging in biology and medicine using ion microscopy., *Microchimica Acta*, **104**, (1-6) January 1991, doi: 10.1007/BF01245536
- [27] J. Dubois, G. Sando, E. N. Lewis, Near-Infrared Chemical Imaging, A Valuable Tool for the Pharmaceutical Industry, G.I.T. Laboratory Journal Europe, No.1-2, 2007.
- [28] Applications of Novel Techniques to Health Foods, Medical and Agricultural Biotechnology.(June 2004).,I. C. Baianu, P. R. Lozano, V. I. Prisecaru and H. C. Lin q-bio/0406047 (<http://arxiv.org/abs/q-bio/0406047>)
- [29] Chemical Imaging Without Dyeing (<http://witec.de/en/download/Raman/ImagingMicroscopy04.pdf>)
- [30] C.L. Evans and X.S. Xie.2008. Coherent Anti-Stokes Raman Scattering Microscopy: Chemical Imaging for Biology and Medicine., doi:10.1146/annurev.anchem.1.031207.112754 *Annual Review of Analytical Chemistry*, **1**: 883-909.
- [31] Eigen, M., Rigler, R. Sorting single molecules: application to diagnostics and evolutionary biotechnology,(1994) *Proc. Natl. Acad. Sci. USA*, 91,5740-5747.
- [32] Rigler, M. Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology,(1995) *J. Biotechnol.*, 41,177-186.
- [33] Rigler R. and Widengren J. (1990). Ultrasensitive detection of single molecules by fluorescence correlation spectroscopy, *BioScience* (Ed. Klinge & Owman) p.180.

- [34] Oehlenschläger F., Schwille P. and Eigen M. (1996). Detection of HIV-1 RNA by nucleic acid sequence-based amplification combined with fluorescence correlation spectroscopy, *Proc. Natl. Acad. Sci. USA* **93**:1281.
- [35] Bagatolli, L.A., and Gratton, E. (2000). Two-photon fluorescence microscopy of coexisting lipid domains in giant unilamellar vesicles of binary phospholipid mixtures. *Biophys J.*, 78:290-305.
- [36] Schwille, P., Haupts, U., Maiti, S., and Webb, W. (1999). Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophysical Journal*, 77(10):2251-2265.
- [37] FRET description (<http://dwb.unl.edu/Teacher/NSF/C08/C08Links/pps99.cryst.bbk.ac.uk/projects/gmocz/fret.htm>)
- [38] doi:10.1016/S0959-440X(00)00190-1 ([http://dx.doi.org/10.1016/S0959-440X\(00\)00190-1](http://dx.doi.org/10.1016/S0959-440X(00)00190-1)) Recent advances in FRET: distance determination in protein–DNA complexes. *Current Opinion in Structural Biology* **2001**, 11(2), 201-207
- [39] <http://www.fretimaging.org/mcnamaraintro.html> FRET imaging introduction

Interactomics

Interactomics is a discipline at the intersection of bioinformatics and biology that deals with studying both the interactions and the consequences of those interactions between and among proteins, and other molecules within a cell^[1]. The network of all such interactions is called the Interactome. Interactomics thus aims to compare such networks of interactions (i.e., interactomes) between and within species in order to find how the traits of such networks are either preserved or varied. From a mathematical, or mathematical biology viewpoint an interactome network is a graph or a category representing the most important interactions pertinent to the normal physiological functions of a cell or organism.

Interactomics is an example of "top-down" systems biology, which takes an overhead, as well as overall, view of a biosystem or organism. Large sets of genome-wide and proteomic data are collected, and correlations between different molecules are inferred. From the data new hypotheses are formulated about feedbacks between these molecules. These hypotheses can then be tested by new experiments^[2].

Through the study of the interaction of all of the molecules in a cell the field looks to gain a deeper understanding of genome function and evolution than just examining an individual genome in isolation^[1]. Interactomics goes beyond cellular proteomics in that it not only attempts to characterize the interaction between proteins, but between all molecules in the cell.

Methods of interactomics

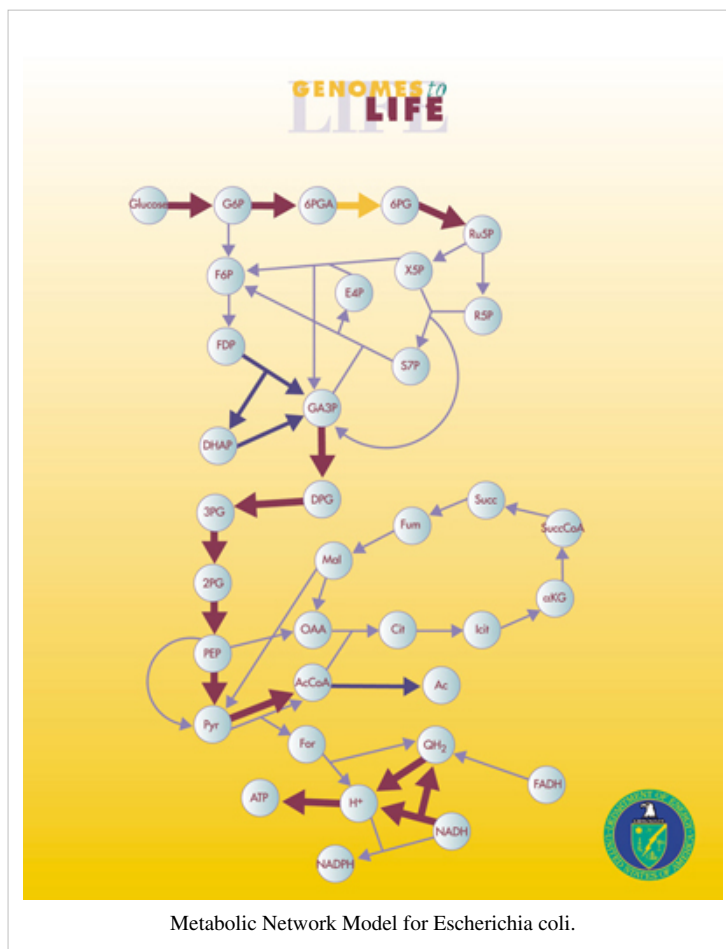
The study of the interactome requires the collection of large amounts of data by way of high throughput experiments. Through these experiments a large number of data points are collected from a single organism under a small number of perturbations^[2] These experiments include:

- Two-hybrid screening
- Tandem Affinity Purification
- X-ray tomography
- Optical fluorescence microscopy

Recent developments

The field of interactomics is currently rapidly expanding and developing. While no biological interactomes have been fully characterized. Over 90% of proteins in *Saccharomyces cerevisiae* have been screened and their interactions characterized, making it the first interactome to be nearly fully specified^[3].

Also there have been recent systematic attempts to explore the human interactome^[1] and^[4].



Other species whose interactomes have been studied in some detail include *Caenorhabditis elegans* and *Drosophila melanogaster*.

Criticisms and concerns

Kiemer and Cesareni^[1] raise the following concerns with the current state of the field:

- The experimental procedures associated with the field are error prone leading to "noisy results". This leads to 30% of all reported interactions being artifacts. In fact, two groups using the same techniques on the same organism found less than 30% interactions in common.
- Techniques may be biased, i.e. the technique determines which interactions are found.
- Interactomes are not nearly complete with perhaps the exception of *S. cerevisiae*.
- While genomes are stable, interactomes may vary between tissues and developmental stages.
- Genomics compares amino acids, and nucleotides which are in a sense unchangeable, but interactomics compares proteins and other molecules which are subject to mutation and evolution.
- It is difficult to match evolutionarily related proteins in distantly related species.

See also

- Interaction network
- Proteomics
- Metabolic network
- Metabolic network modelling
- Metabolic pathway
- Genomics
- Mathematical biology
- Systems biology

External links

- Interactomics.org^[5]. A dedicated interactomics web site operated under BioLicense.
- Interactome.org^[6]. An interactome wiki site.
- PSIBase^[7] Structural Interactome Map of all Proteins.
- Omics.org^[8]. An omics portal site that is openfree (under BioLicense)
- Genomics.org^[9]. A Genomics wiki site.
- Comparative Interactomics analysis of protein family interaction networks using PSIMAP (protein structural interactome map)^[10]
- Interaction interfaces in proteins via the Voronoi diagram of atoms^[11]
- Using convex hulls to extract interaction interfaces from known structures. Panos Dafas, Dan Bolser, Jacek Gomoluch, Jong Park, and Michael Schroeder. *Bioinformatics* 2004 20: 1486-1490.
- PSIBase: a database of Protein Structural Interactome map (PSIMAP). Sungsam Gong, Giseok Yoon, Insoo Jang *Bioinformatics* 2005.
- Mapping Protein Family Interactions : Intramolecular and Intermolecular Protein Family Interaction Repertoires in the PDB and Yeast, Jong Park, Michael Lappe & Sarah A. Teichmann, J.M.B (2001).
- Semantic Systems Biology^[12]

References

- [1] Kiemer, L; G Cesareni (2007). "Comparative interactomics: comparing apples and pears?". *TRENDS in Biochemistry* **25**: 448–454. doi:10.1016/j.tibtech.2007.08.002.
- [2] Bruggeman, F J; H V Westerhoff (2006). "The nature of systems biology". *TRENDS in Microbiology* **15**: 45–50. doi:10.1016/j.tim.2006.11.003.
- [3] Krogan, NJ; et al. (2006). "Global landscape of protein complexes in the yeast *Saccharomyces Cerevisiae* ". *Nature* **440**: 637–643. doi:10.1038/nature04670.
- [4] further citation needed
- [5] <http://interactomics.org>
- [6] <http://interactome.org>
- [7] <http://psibase.kobic.re.kr>
- [8] <http://omics.org>
- [9] <http://genomics.org>
- [10] <http://bioinformatics.oxfordjournals.org/cgi/content/full/21/15/3234>
- [11] http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6TYR-4KXVD30-2&_user=10&_coverDate=11%2F30%2F2006&_rdoc=1&_fmt=&_orig=search&_sort=d&view=c&_acct=C000050221&_version=1&_urlVersion=0&_userid=10&md5=8361bf3fe7834b4642cdda3b979de8bb
- [12] <http://www.semantic-systems-biology.org>

Article Sources and Contributors

Spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=353906901> *Contributors:* 209.234.79.xxx, 5 albert square, ACrush, Aadal, Abdul wali98, AdjustShift, Adoniscik, Afrine, Ahnerpk111, Ahoerstemeier, Akhram, Al-khowarizmi, Alansohn, Andre Engels, Andycjp, Angrysockhop, Annabel, Aspiring chemist, Atlant, Austin512, Azo bob, Bci2, Beano, Becritical, BenFrantzDale, Bensaccount, Bj norge, Borgx, Bryan Derksen, Camw, CanaryMBurns, Charles Gaudette, Charles Matthews, Chemistry isMyLife, Christian List, Coffea, CommonsDelinker, Conversion script, Cortonin, Cremepuff222, Ctro936, Curps, DMacks, Danski14, Deglr6328, Dekisugi, Dfbaum, Dicklyon, Dkroll2, Dpoduval, Dr-b-m, DrBob, Duncanssmith, Dysprosia, Edgar181, El C, Eleassar, ErikvDijk, Erwinrossen, Fastfission, Femto, Flewis, FocalPoint, Fratrep, Fredericks, Fredrik, Frikensmurf, Gentgeen, Gerkleplex, Giflite, Goldenrowley, Guillermo con sus ruedas, Gurch, H Padleckas, Hall Monitor, Hankwang, Harbir93, Harris7, Headbomb, Heron, Holdendp, Icairns, Ignoramibus, Ikanreed, IronGargoyle, J2hawiki, JNW, Jaeger5432, Jameswbk, Javert, Jcwf, JeremyA, JohnCD, Ken6en, Kevyn, Kilva, Kkmurray, Kopeliovich, Kraftlos, La Pianista, LeaveSleaves, Logger9, Marek69, Marj Tiefert, Martin Hedegaard, Martyjmch, Mav, Mejor Los Indios, Melchoir, Michael Hardy, Mihano, Mikeo, Modlab, Mrs Trellis, Mtk180, Mts0405, Nbkoneru, Nicktaylor100, North Shoreman, Northryde, Nuno Tavares, Oxymoron83, PDManc, Pcarbonn, Pdcook, Peter, Petergans, Peterlewis, Pharmacomancer, Pit, Pizza Puzzle, Pjvpjv, Prgo, Prolog, Pyrospirit, Qxz, R'n'B, RG2, Radagast83, Ratsbew, Reddi, Renxa, Retired username, Rjwilmsi, Rd200, RoadieRich, Rob Hooft, Rocoptics, Ronningt, RoyBoy, Rubin joseph 10, Rvoorhees, S4wilson, Safalra, Sam Hovevar, Sbialkow, Scg, SemperBlotto, Shadowjams, Sharare, Shell Kinney, Shovkowska, Smack, Snags, Sodium, Srleffler, Srnee, StefanP1, Stone, StradivariusTV, T-borg, Taxman, Teges17, TheIntersect, Thecurran91, Tibor Horvath, Tide rolls, Tim Starling, Tsemii, Tukan, Voyajer, Vsmith, Welsh, WiKi, Wik, Will.i.am, Williamheyn, WinterSpw, Wk muriithi, Wpostma, Zereschk, Zotel, Zroutik, 304 anonymous edits

Fourier transform spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=353620276> *Contributors:* AJim, Asterion, Bci2, Bersbers, Berserkerus, BigFatBuddha, Bobby1011, Christopherlin, Conversion script, Damian Yerrick, Deglr6328, E104421, EndersJ, Eprb123, Graeme Bartlett, Graham87, Greggklein, Guillom, Hankwang, Harold f, Haydarkustu, HelgeStenstrom, Jaraalbe, Jcwf, John.lindner, Jonathan F, Kcordina, Kingpin13, Kkmurray, Martyjmch, Matanzb, Michael Hardy, Nikai, Nitrogen15, PBSurf, Peter, Peterlewis, Publicly Visible, Rifleman 82, Rnt20, Ronningt, Roybb95, Shyrnes321, Seidenstud, Skier Dude, Slapidus, Smeyer1, Stannered, Stokerm, Sverdrup, Taw, Thinkinnng, Tim Starling, Veinor, Victorsong, Vsmith, 74 anonymous edits

Quantum mechanics *Source:* <http://en.wikipedia.org/w/index.php?oldid=355114185> *Contributors:* 06twalke, 100110100, 11341134a, 11341134b, 128.100.60.xxx, 172.133.159.xxx, 21655, 3peasants, 4RM0, 5Q5, 64.180.242.xxx, 65.24.178.xxx, A. di M., APH, Aaron Einstein, Abce2, Academic Challenger, Acalamari, Acroterion, AdamJacobMuller, Addionne, Addshore, AdevarTruth, Adhalanay, AdjustShift, Adolfman, Adventurer, Agge1000, Ahoerstemeier, Aitias, Ajcheema, Akriasas, Akubra, Alai, Alansohn, Ale jrb, Alejo2083, AlexBG72, AlexiusHoratius, Alfio, Alison, Amareto2, Amarvc, Amatulich, AmiDaniel, Amit Moscovich, Ancheta Wis, AndonicO, Andrej.westermann, Andris, AndriyK, Andy chase, Angelsages, AnnaFrance, AnonGuy, Anonwhymus, Antandrus, Antixt, Antonio Lopez, Anville, Ap, ApolloCreed, Apparition11, Arjen Dijksmann, Arjun01, Army1987, Arnero, Arthur chos, Asdf1990, Ashujo, Asmackey, AstroNomer, Asyndent, AtticusX, AugustinMa, Austin Maxwell, Awolf002, AxelBoldt, Axlloren, BM, BMD2101, Badgettrg, Baker APS, Barak Sh, Barbara Shack, Bart133, Batmanand, Bcasterline, Bci2, Bccor, Bdesham, Beano, Beatnik8983, Beek man, Belsazar, Bensaccount, Bento00, Benvirg89, Benvogel, Beta Orionis, Betacommand, Betterusername, Bevo, Bfiene, Billosbyislonelypart2, Bkalfat, BlastOButter42, Bmjui, Bobo192, Bongwarrior, Bookaling, Borki0, Bowfolknowledge, Brettwats, Brews ohare, Brina700, Broughm96, Bsharyv, BuffaloBill90, Bustamonte2003, Bytgenwulf, C.c. hopper, CALR, CBMIBM, Clreland, CSTAR, CSfar, CYD, Caiaffa, Ccin47, Can't sleep, clown will eat me, CanadianLinuxUser, Canderson7, Capricorn42, Captain Quirk, Captain-turck, Carborn1, CardinalDan, Caroline Thompson, Carowinds, Casomerville, Catgut, Centic, Chairman S., Chao129, Chaos, CharlotteWebb, Charvest, Chaujie328, Chenyu, Chetvorno, ChooseAnother, Chris 73, Chris5858, Christian List, ChristopherWillis, Chun-hian, CiA10386, CieloEstrellado, Cigarettezz, Cipapuc, Civil Engineer111, Clarinace63, Clarkbhm, Cleric12121, Closedmouth, ClovisPt, Cmicahel, Comestyles, Commander zander, Complexica, Connormach, Conversion script, CosineKitty, Couki, Cp111, Cpcheung, Cpl Syx, Creativethought20, Crowsnest, CryptoDerk, Curlymeatball38, Cybercobra, Cyclonenim, Cyp, DJIndica, DL5MDA, DV8 2XL, DVdm, Da monster under your bed, Dale Ritter, Dan Gluck, Dan100, DangApricot, DanielGlazer, Dannideak, Darguz Parsilvan, DarkFalls, Darrellpenta, Darth Panda, Das Nerd, Dataphile, Dauto, David Gerard, David R. Ingham, David Woolley, Dbroesch, DeCooper, DeFoaBuSe, DeadEyeArrow, Decumanus, Deldot, DenisDiderot, DerHexer, Derek Ross, Derval Sloan, DetlevSchm, Diafanakrina, Dickwet89, Dingoatscritch, Discospinster, Djr32, Dmblub, Dmr2, Doc Quintana, Dod1, Dominic, Donbert, Donvinzk, Dpotter, Dragon's Blood, Dreadstar, Dreftymac, Drini, Dspradau, Dtgim, Dylan Lake, E Wing, EMC125, Ed Poor, Edsanville, Edward Z. Yang, El C, El Chupacabra, ElectronicsEnthusiast, Ellenseel, Elitedarklord dragonslyer 3.14159, Ellmist, Eloquence, Enochlau, Enormousdude, Enviroboy, Eprb123, Eric-Wester, Eurobas, Evan Robidoux, Eveross, Ex nihil, FaTony, Factosis, FagChops, Falcorian, Farosdaughter, FastLizard4, Fastfission, Fearingfearitself, Feezo, Feinoha, Femto, Figma, Finalnight, Firas@user, Firewheel, Firien, Floorsheim, FloridaSpaceCowboy, Foobar, Foobar, Four Dog Night, FrancoGG, Frankenpuppy, Freakofnuture, Fredkinfollower, Fredrik, Freemarker, FreplySpang, Frosted14, Fvww, GLaDOS, GTubio, GaaraMsg, Gabbe, GabrielF, Gaius Cornelius, Gandalf61, Gary King, Gcad92, Geek84, Geometry guy, GeorgeLouis, GeorgeOrr, Gierens27, Gierens22, Giflite, Givegains, Gjd001, Glennwells, Gnixon, Gogo Dodo, Gpgarrettboast, G3rk3 b0i, Grafnita, Graham87, GrahamHarzo, GreenReflections, Greg HWWOK Shaw, GregorB, Grey Shadow, Grim23, Grunt, Guitarspecs, Guy82914, Guybrush, Gwernol, Gyan, H.W. Clihor, H0riz0n, H1nomaru senshi, Hadal, HaeB, HamburgerRadio, Han Solar de Harmonics, Handsome Pete, HappyCamper, Harp, Harriemkali, Harryboyles, Harrytipper, Hdeasy, Headbomk, Hekoshi, Henry Delfom, Heptarchy of teh Anglo-Saxons, baby, Herakles01, Hermione1980, Hidaspal, Hikui87, Hobartimus, Holy Ganga, HounsGut, Hughgr, Huon, Husond, Hut 8.5, Hviltvis, II MusLiM HyBrRiD II, IJMacD, IRP, IW.HG, Ilyushka88, Imalad, Info D, Iqesurnioe, Iridescent, Islander, Itub, Ivan Štambuk, Ixf64, Izodman2012, J.delanoy, JForget, JMK, JSpudeman, JSpung, Jack O'Lantern, Jackthegrape, Jaganath, Jagbag2, Jagged 85, Jake Wartenberg, JakeVortex, Jaleho, Jan eissfeldt, Jaquesthehunter, Jason Quinn, Javafreakin, JayJasper, Jebba, Jecar, Jeffrey Vest, JinJian, Jitse Niesen, Jmartinsson, Jni, JoJan, JockK, Joe hill, JohnCD, Johnleemk, Johnstone, Jon Cates, Jordan M. JorvisS, Joseph Solis in Australia, JoshC306, Joshuaflimer, Joss, Jostylr, Joustler, Juliancolton, Joshuaflax, Jz, K-UNIT, KCinCD, KHAMsun, KTC, Kadski, Kaesle, Kan8eDie, Karol Langner, Katalaveno, Katzmik, Kding, Ke4roh, Keenan Pepper, Kenneth Dawson, Kesaloma, Ketsuekigata, Kevin aylward, Kevmitch, KeyStroke, Khatharr, Kingdon, KnowledgeOfSelf, Knutux, Koepfinger, Krackenback, Krazywrath, Kroflin, Kriar, Kyokpae, Kzollman, L-H, LWF, La hapalo, Lacrimosus, Lagagal, Latka, Laurascudder, Lavateraguy, Le fantome de l'opera, Lear's Fool, Leavage, Leesy1106, LeoNomis, Leoeadec, Lethe, LiRoss2009, Light current, Lightmouse, Ligulem, LilHelpa, Linas, Lindsavi, Linuxerick, Lithium412, Lmp883, LokiClock, Lontax, Looie496, Loom91, Loopism, Looxix, Lotje, Lowellian, LucianLachance, Luk, Luna Santin, Lupo, MBlue2020, MER-C, MK8, MKoltnow, MPS, Mackaf92, Macy, Mafiaparty303, Magnus, Magog the Ogre, Mahommed alpac, Marek69, Mark Swiggle, Marks87, MarphyBlack, Martial75, Martin Hedegaard, Matt McIrvn, Matb112885, Matthewasse1, Matty j, Maurice Carbonaro, Mav, Maximillion Pegasus, Mbell, McSly, Mcdutchie, Mckaysalishbury, Mebden, Megankerr, Mejor Los Indios, MelicansMatkin, Menchi, Meno25, Metaldev, M giganteus1, Michael C Price, Michael Courtney, Michael Hardy, Michael9422, MichaelMaggs, MichaelVernonDavis, Michele123, Midnight Madness, Midom, Mifter, Mिकाेय, Mike Peel, Mike6271, Mikeo, Minimac's Clone, Mitchellslims08, Mjamja, Mjb, Mtk2rhino, MoForce, Modusoperandi, Moeron, Moink, Mojska, Mollmop, Moncief, Monedula, Moogoo, Moskvax, Mpatel, MrJones, Mrsastrochicken, Mrvanmer, MtB, Muizj, Mustbe, Mxn, Myasuda, Mygerardormance, N1RK4UDSK714, Nabla, Nagy, Nanobug, Nathan Hamblen, Naturelles, NawlinWiki, NellieBly, Netalarm, Neverquick, NewEnglandYankee, Nibbleboob, Nick Mks, Nick125, NickBush24, Nielad, NightFalcon90909, NijaMunki, Nikai, Nin0rz4u 2nv, Niout, Niz, NoEdward, Noisy, Norm, North Shoreman, NotAnonymou0, Nsaa, Nturtion, NuclearWarfare, Numb3, Nunquam Dormio, Nv8200p, Nxtid, Octavianvs, Odenluna, Oleg Alexandrov, Olivier, Ondon, Onesmoothlefty, Onorem, Oo64eva, Optin, OrgasGirl, Otolemur crassicaudatus, Ott, OwenX, Owl, Oxymoron83, P99am, PAR, PGWG, Palica, PandaSaver, Papadope, Pascual, Tesson, Pasky, Pathoschild, Patrick0Moran, Paul August, Paul Drye, Paul E T, Paul bunion, Paulc1001, Paulraine, Pcarbonn, Pchov, Peace love and feminism, Peter, Peterdjones, Petgraveyard, Petri Krohn, Pfalstad, Pgb23, Pkg, Phil179, Philip Trueman, Philippe, Philmurray, Phoenix12, PhysSusie, Phys, PhysicsExplorer, Physika, Physprof, Piano non troppo, Pilotguy, Pippu d'Angelo, Pj.debruin, Pjacobi, Pleasantville, Pokemon274, Polgo, PorkHeart, Positivetruthintent, Possum, Prashanthins, Pxfbird, Quackbumper, Quantumtmotion, Quercus basaeacchicensi, Quuxphulso, Qwertuio 132, Qxz, Qzd800, RG2, RJaguar3, Raaziq, Radzewicz, Rakniz, RandomXYZb, Randy Kryn, Razataza, RazorICE, Rdsmith4, Reconsider the static, Rednblu, Reinis, Rettetast, Rex the first, Rgambie, Rich Farmbrough, Richard L. Peterson, Richard001, Ridge Runner, Rightly, Rnb, Rnpg1014, Roadrunner, Robb37, Robbjedi, RobertG, Robertd514, RobinK, Roebier, Rogper, Romaioi, RonhJones, Rory096, Roxbreak, RoyBoy, Royalbroil, Royalguard11, Rursus, Russel Mccpinigin, RxsS, Ryan ley, RyanB88, SCZenz, SFC9394, SJP, SWAdair, Sadi Carnot, Saibod, Sajendra, Samlyn.josfyn, SanfordEsq, Sango123, Saperaud, Savant1984, Sbandrews, Schumi555, ScienceApologist, Scott3, ScudLee, Sean D Martin, Search4Lancer, Seawolf1111, Seba5618, SeriousGrinz, SeventyThree, Shanemac, Shanes, Sheliak, Shipunits, Shoyer, Signalhead, Silence, Silly rabbit, Silsors, Sionus, Sir Vicious, SirFozzie, Sjoerd visscher, Skewwhiffy, Skizzik, SkyWalker, Slaniel, Smallpond, Smeat, Smyth, SnappingTurtle, Snowolf, Snoyes, Sokane, Sokratesinabasket, Soliloquial, Solipsist, Solo Zone, Someguy1221, Soporatemoon, SpaceMonkey, Spearhead, Specs112, Spellcast, Spiffy sperry, Sriram sh, Srleffler, St33lbird, StaticGull, StefanPernar, Stentor7, Stephan Leclercq, Stephen e nelson, Steve Quinn, Stevenganzburg, Stevenj, Stevenwagner, Stvertigo, Stiangk, Stone, StonedChipmunk, Storm Rider, Stormie, StradivariusTV, Sukael, Sukaj, SunDragon34, Superior IQ Genesis, Superlions123, Superstring, Suppersmart, Swegei, Swwright, Syrrhiss, TTGL, Ta bu shi da yu, TakuyaMurata, Tangotango, Tantalate, Taral, Tarotcards, Taylo9487, Tayste, Tcnv, Teardrop onthefire, Template namespace initialisation script, Terra Xin, Texture, The Anome, The Captain, The Firewall, The High Fin Sperm Whale, The Mad Genius, The Red, The Spam-a-nata, The Thing That Should Not Be, The ed17, The wub, TheKoG, Theresa knott, Things, Thunderhead, Tide rolls, Tim Shuba, Tim Starling, TimVickers, Timwi, Titoxd, Tirkman, Tohd8BohaiithuGh1, Toofy mcjack34, TripLikeIdo, Tromer, Trusilver, Truthnlove, Ttasterul, Tycho, TylerM37, Tylenm17, UberCryxic, Ujm, Ultratomo, Uncle Dick, Uptidatelper, Urbansky, Utcursch, V81, Van helsing, VandalCruncher, Vb, Velvetron, Venu62, Vera.tetrix, Versus22, Viduoke, Vitalikkk, Vlatkovedral, Voyajer, Vsmith, Vudujava, WAS 4.250, WmdeMuyncn, Wafulz, Wagggers, WarEagleTH, Wareagles18, Wavelength, Waxigloo, Wayward, Weiguxp, Welsh-Jord-Lowri, Werdna, WereSpielChequers, Weregerbil, Wik, WikHed, Willieru18, WilliumWeales, Willking1979, Willtron, Wimt, Winston365, Wk muriithi, Wknight94, Woohookitty, Word2need, Wquester, X41, XJAmRastafire, XTRENCHARD29x, Xana's Servant, XeniaKon, Xezlec, Xihr, Xy7, Yamaguchi先生, Yath, Yeahyeahkickball, Yellowdesk, Yokabozeez, Zakuragi, Zarniwoot, Zenbullets, Zero over zero, Zhinz, Zigger, ZiggyZig, Zolot, Zondor, Zooloo, Zslevi, Александр, 1787 anonymous edits

Quantum field theory *Source:* <http://en.wikipedia.org/w/index.php?oldid=353812218> *Contributors:* 130.64.83.xxx, 171.64.58.xxx, 9.86, APH, Acalamari, Ahoerstemeier, Albertod4, Alfred Centauri, Alison, AmarChandra, Amareto2, Amarvc, Ancheta Wis, Andrea Allais, Archelon, Arnero, AstroPig7, Bambaiah, Bananan, Banus, Bci2, BenRG, Bevo, Bkalfatuf, Bkennis2006, Bckavnger, Bob108, Breds ohare, Bt414, CALR, CYD, CambridgeBayWeather, Caracolillo, Charles Matthews, Ciphers, Complexica, ConCompS, ConradPino, Conversion script, Custos0, Bcyerobra, Cyp, D6, DJIndica, Dan Gluck, Dauto, DefLog, Dj thegreat, Dmkhowarth26, Docboato, Drschwarz, Dylan1946, Egg, El C, Erik J, Etale, Four Dog Night, Gandalf61, Garion96, Gcbirzan, GeorgeLouis, Giflite, Gjseejith, Glenn, Graham87, GrahamHardy, HEL, Hairy Dude, Hanish.polavarapu, Headbomb, Helix84, Henry Delfom, HexaChord, HowardFrampton, Ht686rg90, Hyqeom, IRP, IZAK, Igodard, Igorivanov, Iridescent, Itinerant1, J.delanoy, JEH, Jamie Lokier, Jecht (Final Fantasy X), Jeepday, Jim.belk, Jmnbatista, Joseph Solis in Australia, Juliancolton, KarlHallowell, Karl Langner, Kasimann, Kenneth Dawson, Knowandgive, Kromatol, LLHolm, Leapold, Lerman, Lethe, Likebox, Looxix, Looxix, Maliz, MarSch, Marie Poise, Markisgreen, Marks, Martin Kostner, Masudr, Mav, Mbell, Mboverload, Mclid, Melchoir, Mendicus, Mentifisto, M giganteus1, Michael C Price, Michael Hardy, Moltrix, Mpatel, Msebast,

MuDavid, N shaji, Neparis, Newt Winkler, Nikolas Karalis, Niout, Northgrove, Nvj, Odddmonster, Opie, PSimeon, Palpher, Paolo.dL, Pcarbonn, Phys, Physicistjedi, Pinkgoanna, Pjacobi, Pt, Puksik, QFT, R.ductor, R.e.b., RE, RG2, Raphtee, Rebooted, Roadrunner, Robert L, RogueNinja, Rotem Dan, Rudminjd, RuudVisser, SCZenz, SebastianHelm, Semperf, Shanes, Sheliak, Shivav, Sietse Snel, SirFozzie, Skou, SirPellazie, Sreffler, Stephan Schneider, Stevertigo, StewartMH, Tamtamar, Template namespace initialisation script, The 80s chick, The Anome, The Original Wildbear, The Wild West guy, Thialfi, Thingg, Threepounds, Tim Starling, TimNelson, Timwi, Tlabshier, Tobias Bergemann, Tom Loughheed, Truthnlve, Tugjob, UkPaolo, Urvabara, Van helsing, Vanderdecken, Varuna, Victor Eremita, Wavelength, Wik, Witten Is God, Wwheaton, XJAmRastafire, Yndurain, Zarniwoot, 254 anonymous edits

Algebraic quantum field theory *Source:* http://en.wikipedia.org/w/index.php?oldid=24316506 *Contributors:* Charles Matthews, Chtito, Conscious, CryptoDerk, DefLog, Gauge, Geometry guy, Giftlite, Jag123, Kate, Lethe, MFH, Michael Hardy, Phys, Rich Farmbrough, Rorro, West Brom 4ever, 9 anonymous edits

Local quantum field theory *Source:* http://en.wikipedia.org/w/index.php?oldid=328904570 *Contributors:* Charles Matthews, Chtito, Conscious, CryptoDerk, DefLog, Gauge, Geometry guy, Giftlite, Jag123, Kate, Lethe, MFH, Michael Hardy, Phys, Rich Farmbrough, Rorro, West Brom 4ever, 9 anonymous edits

Algebraic logic *Source:* http://en.wikipedia.org/w/index.php?oldid=340372691 *Contributors:* CBM, CBM2, Chalst, Charles Matthews, Elwikipedista, EmiIJ, Giftlite, Gregbard, Hmains, Jitse Niesen, Koavf, Mhss, Michael Hardy, PWilkinson, Palnot, Strangename, The Tetrast, Trovatore, 11 anonymous edits

Quantum logic *Source:* http://en.wikipedia.org/w/index.php?oldid=353702379 *Contributors:* Alvin Seville, Andris, Angela, Archelon, Argumzio, Aster Rainbow, BD2412, Bci2, CSTAR, Charles Matthews, Cybercobra, DJIndica, David edwards, Dcoetzee, Dmr2, Dysprosia, Edward, Epsilon0, GTBacchus, Gaius Cornelius, Gene Ward Smith, Giftlite, GordonRoss, Gregbard, Hairy Dude, Headbomb, Icairns, Ilan770, Jengod, John Baez, Kimberlylg, KnightRider, Kuratowski's Ghost, Kzollman, Lethe, Linas, Lucidish, Mct mht, Meznaric, Mhss, Michael Hardy, Modify, Oerjan, Parkyere, Pohta ce-am pohtit, PowerUserPCDude, Rsimmonds01, Sheliak, Shlomi Hillel, StevenJohnston, Stevertigo, T=0, Trovatore, Uncia, V79, Zumbo, ٢٤٤٣, 40 anonymous edits

Quantum computer *Source:* http://en.wikipedia.org/w/index.php?oldid=354261630 *Contributors:* -Ozone-, 194.117.133.xxx, Imujin22, 4lex, A5b, AAAAA, AWeishaupt, Aarchiba, Abdullahazzam, Ajb, Albedo, Ale2006, Alex R S, Alksentrs, Alsandro, Amcfreely, Amwebb, AndrewStuckey, Andrewpmk, Andris, Antandrus, Antlegs, Anville, Anwar saadat, Arcfrk, Archelon, Arnero, ArnoldReinhold, Asmeurer, Asparagus, Atropos235, Auric, AxelBoldt, B9 hummingbird hovering, Bci2, Beagel, Beeban, Ben Standeven, Bencoder, Bender235, Bevo, Bibliosapien, Bigmantonydz, Biheo, Bjohnsn00, Blotchley, Bobby D. Bryant, Bobo192, Bond4154, Booyabazooka, Brandon, Brighterorange, Brilliand, Bryan Derksen, Bubba hotep, BumbleFootClown, Burgaz, CBM, CSTAR, CYD, Captanpluto123, Card, Cenarium, Charles Matthews, CharlesGillingham, Chris 73, CloudNine, Comestyles, Connelly, Couchpotato99, Count Caspian, Craig Pemberton, Creidieki, Crunkcar, Cryptonaut, Curtdbz, Cyan, Cyp, DV8 2XL, Dan East, Dan Granahan, Danko Georgiev MD, Danny, Dar-Ape, DarkFalls, DarkShroom, Daveapp, David Battle, David Gerard, David n m bond, DavidB, DavidLevinson, Davidgothberg, Davidmznz, Dead3y3, Dethme0w, Dgrant, Dr.K., DuckFeather, Durval, Dysprosia, Eaglizar, Edd Porter, Edonovan, Edward, Eigenlambda, Ekangas, El C, Elb2000, Eluchil, Epo, ErinM, Eurobas, Eyu100, FKmailiW, Fattony77, Favonian, Fgrosshans, Foxxygirltamara, Frankenpuppy, Fullstop, Furrykef, Fyyer, G Colyer, GBL, GaaraMsg, GaeusOctavius, Gail, Gaius Cornelius, Galwhaa, Gary King, Gauge, Gentgeen, George100, Giftlite, Gioto, Glenn, Gnixon, Gnosmitharmy, Graham87, Green caterpillar, GreenoRB, Gtq207u, HRV, Hadal, Hanjabba, Hannover.post, Hans Adler, Harold f, Harry Potter, Harryboyles, Hatethedj, Havanafreestone, Headbomb, Hfatedge, Hmrox, Hulagutten, HumphreyW, Husond, Hyandat, Hyperdivision, Iain David Stewart, Ihope127, Ikh, Iliia Kr., Illspirit, Indosauros, Insurrectionist, Intgr, Inwind, Isaacsurh, Iseeaboar, Isnow, Ispy1981, J.delanoy, JMK, JRGregory, Jao, Jarekadam, JavOs, Jbw2, Jeargle, Jeff G., Jengod, Jitse Niesen, Jij137, Jksad, Jni, JohnCD, Jon513, JonHarder, Jorfer, Jotomicron, Jrockley, Justin Stafford, Jwoehr, KD5TVI, Kannan karthik, Kate, Keegan, Keenan Pepper, Kevin B12, Kevyn, Kine, Kizeral, Korepin, Ksn, Kuru, LC, Liftarn, Ligulem, Linas, Linus M., Looxix, Lostart, Lotu, LunatikOwl, Lunkwill, Lupin, MBisanz, MER-C, MagnaMopus, Manoliisfat, Marco de Mol, Marksuppes, Martaen, MartinSieg, Mat8989, Matt Crypto, Mattblack82, Matthewsism, Maurice Carbonaro, Mav, Maximus Rex, MementoVivere, Michael Hardy, MichaelMcGuffin, Michal Jurosz, Mike1024, Millerc, Mindmatrix, Miym, Mjager, Moment, Moonriddengirl, Mpassman, MrOllie, Mro, Muj0, MuncherOfSpleens, Nidickson, Nol888, Noldoaran, Norm mit, Not enough names left, NotThatJamesBrown, Nurg, Nzv8fan, Octopus-Hands, Officiallover, Ojigiri, Oleg Alexandrov, Oliver Pereira, Onaillim, Oneyd, Onorem, Pagrashtak, Pak21, Pakaran, Palfrey, PatrickHoo, PaulCook, Peter bertok, Peterdjones, Phil Boswell, Phoneused, Physis, Piano non troppo, Pletsche, Pohta ce-am pohtit, Poor Yorick, Populus, Powo, Pwjb, Pyrosiprit, Qwertyus, REQ2, RG2, RHaworth, RJRocket53, RTC, Racklever, Rajrajmarley, Ravedave, ReallyMale, Red Thunder, RedWolf, Remy B, Revived, Rich Farmbrough, Ripper234, Roadrunner, Robert Merkel, RobertG, RobinK, Robma, RonanSandford, Ronz, RoyBoy, Ruud Koot, Rvwww, Sajendra, Sam Hocevar, Sambarrows, Sanders muc, SandyGeorgia, Sathimantha, Scott McNay, Scottcraig, Selain03, Seraph 31, Shepelyansky, Shoyer, Simetrical, SimonMayer, Skippydo, Sligocki, Smite-Meister, Snakeyes (usurped), Snowolf, Speedplane, Spikeuk14, Spin-Half, Sploo22, Squiggle, Steamturn, Stevertigo, Ssang, Supaman89, Superm401, Susvolans, Svick, Symmetrysinger, TALlama, Taed, Tajo, Talon Artaine, Tarotcards, Tboonsun, Techman224, Tellyaddict, That Guy, From That Show!, The Belgain, The Rogue Penguin, Tim Starling, Timwi, Tobcy75, Togo, Tom harrison, Tomatoman, Tomstdenis, Tre1234, Tree Biting Conspiracy, Tricky Wiki44, Troelsj, Turdus, Turnstep, Ultra megatron, Uncle G, UncleDouggie, Undecided, Vaspian, Verbal, Vincenzo.romano, VodkaJazz, Wafulz, Weekwhom, Wereon, Whispering, Who, Wikiborg, Wikiklrsc, Wikiscient, WilliamKF, Wim van Dam, WindOwl, Wish wellingtons, Wolfkeeper, Woohookitty, Wrdavenport, Wtanaka, X1011, Xiong Chiamiov, Yongxiangu, Zflood, 629 anonymous edits

Quantum chemistry *Source:* http://en.wikipedia.org/w/index.php?oldid=354070567 *Contributors:* 144.189.40.xxx, 208.40.185.xxx, 4lex, Acroterion, Alansohn, Alex '05, Auntof6, Ayla, BTDenyer, Bci2, Bduke, Bob, Brian Y, Bubbia, CDN99, Capecodeph, ChemGardener, Chuunen Baka, CloudNine, Cmdrjameson, CommonsDelinker, Conversion script, Cool3, Cypa, Danhe, EdJohnston, Edsanville, EmilyT, Euryalus, Fygoat, Gentgeen, Gershom, Giftlite, Glenn, GregorB, Haljalad, HappyCamper, Headbomb, Holdran, Hugo-cs, Ian Pitchford, Ithacagorges, Itub, Jantop, JerrySteal, Jheald, Kaliumfredrik, Karol Langner, Keenan Pepper, Keilana, Koinut, Krash, La goutte de pluie, Lampuchi, Ligulem, Lijuni, Looxix, M stone, Martin Hedegaard, Meisfunny, Milo, Nickptar, Noisy, Nzzl, Okedem, P99am, Perelaar, RA0808, Ratol, Rifleman 32, SHL-at-Sv, SQL, Sadi Carnot, Salsb, Shalom Yechiel, Shanel, Sidhekin, Smoe, Sunev, Tasudtry, Terhorstj, Timwi, UninvitedCompany, Vb, Vgy7ujm, Vig vimarsh, Voigfdsa, Vsmith, W.F.Galway, Wiki alf, Xebvor, Yurik, Zarniwoot, Zeimusu, Александр, ٢٤٤٣, 144 anonymous edits

Density functional theory *Source:* http://en.wikipedia.org/w/index.php?oldid=353188446 *Contributors:* Ae-a, Agilemolecule, AlbertoCastro, Apple2, AskHL, Azo bob, Bachrach44, Baxxterr, Bduke, BluePlatypus, Bob K31416, Brews ohare, Buriti, Bwschnei, Charles Matthews, Chemuser, Chymicus, Dirac66, Ebuchol, Edsanville, Emersoni, Enpi, Evgeny, FelixP, Fuzheado, Gafnero, Geboy, Geling, Giftlite, Headbomb, Hourahine, Iamthealchemist, Ianbolland, Islanes, Isnow, Itamblyn, JaGa, Jitse Niesen, Joy, Karol Langner, Lfh, Lucaskw, MaxBoldin, Michael Hardy, MuDavid, NNemec, Nick Mks, Ouji-fin, P99am, Petulda, Poszwa, Pt, Raghunathan, RangeK, Rhobite, Roib1, Royalbooksnap, Rundquist, Ruud Koot, Salsb, Samprox, Sbandrews, Sbo, SebastianHelm, Shaddack, Spellchecker, Stone, Svenbor, TDogg310, THEN WHO WAS PHONE?, Tdouno, Terhorstj, Tim Starling, Tobias Bergemann, Toulouse, Trigger hippie77, V8rik, Vexedd, Vmilman, Vsmith, Wavefunk, WijzeWillem, Wik, WikipAcct, WilliamDParker, Wiz9999, Wolf2046, Xavier andrade, Youandme, Zanimum, Zarniwoot, 192 anonymous edits

Birefringence *Source:* http://en.wikipedia.org/w/index.php?oldid=353773007 *Contributors:* A. B., Ackbeet, Avihu, Bluemoose, Brews ohare, Bryan Derksen, Catskul, Chocofever, Ciphers, Cmdrjameson, Conscious, Corvus cornix, Cutler, DKToptics, Dougher, DrBob, Drewmk2, Elert, Ellieadaedanaforever, Ellywa, Foofighter20x, GeoGreg, Hankwang, Headbomb, Heron, Hu12, Icep, JTN, Jeroen94704, JerrySteal, Johnpseudo, Kar.ma, Karol Langner, Keenan Pepper, Kodang, Krlh8, Laundrypowder, Leeworth, LiDaobing, Lzur, Michael Hardy, Mikael Häggström, MrBell, Ms2ger, Mwtoews, NaBUru38, Nvpntatlwyer, ObsessiveMathsFreink, Ojigiri, Paolo.dL, Peterlewis, Professorgt, Pwjb, Qfn247, Quantumobserver, Rikvoerman, Salthebad, Saperaud, Sdornan, Shramper, Siim, Spiegelberg88, Sreffler, Stefan.bucur, Steve Quinn, Tantalate, Twthmoses, Ufim, Vsmith, YoavShapira, Ytterbium, 98 anonymous edits

Polarization spectroscopy *Source:* http://en.wikipedia.org/w/index.php?oldid=336954187 *Contributors:* Evgeny, Jovianeye, Rich257, ZooFari

Polarized IR Spectroscopy *Source:* http://en.wikipedia.org/w/index.php?oldid=16987408 *Contributors:* 194.200.130.xxx, Aboalbiss, Ahoerstemeier, Anna Lincoln, Annabel, Antandrus, Arcadian, Arnero, Ary29, Bensaccount, BigFatBuddha, Biophysik, Bobthebuilder37, Bomemir, Borgx, BountyTJ, Bubba hotep, CLW, Calaschysm, Charles Matthews, Christopherlin, Chuck Sirloin, Cobi, Coffee, CommonsDelinker, Conversion script, Cwkmall, DMacks, DavidRKelly, Deglr6328, DerHexer, Dieter Baurecht, Dr.Soft, Drbreznjev, Drmies, El C, Eno-ja, Fieldday-sunday, Finalnight, Francs2000, Freestyle-69, Fresheneesz, Fuhghettaboutit, Fyver528, Gentgeen, GeorgHH, GermanX, Giftlite, Gilliam, Greggklein, Grimlock, Guillom, HYPN2457, Hankwang, HappyCamper, Haukurth, Heron, Hesacon, Hollgor, II MusLiM HyBRiD II, Ian Pitchford, Imedio, Informationtheory, J.delanoy, Jackol, Jaraalbe, Jcwif, Johnbrownsbody, Junglectat, Jusdafax, Kcordinia, Kkmurray, Kwiskool, Lifer21, Lightmouse, Littleghoti, Logger9, Lorenzarius, LouisBB, LukeSurl, Martyjmch, Materialsscientist, Michael Hardy, Mjwlans, Mythealias, Nakon, NewEnglandYankee, Nivix, Nmawth, Old Moonraker, Peterlewis, Pharmacomancer, Pit, Punctilius, Quadell, Quantockgoblin, Rifleman 82, Rob Hoof, Ronningt, SABenyunes, Sam Hocevar, SantosH5, Shalom Yechiel, SK9am, Smokefoot, Someguy1221, Srnec, Stephenb, Stokern, SuperTycoon, TechPurism, The wub, Tiago Becerra Paolini, Urbansky, V8rik, Vcelloho, Vector Potential, Vegaswikian, Veinor, Visor, Vsmith, Werson, Wikieditor06, Wmahan, Yashkochar, 253 anonymous edits

Circular dichroism *Source:* http://en.wikipedia.org/w/index.php?oldid=355135441 *Contributors:* AJim, Andrew Rodland, Atlant, Bci2, Bensaccount, Biophysik, Bjsamelsonjones, Bryan Derksen, ChemGardener, Christopherlin, Crystal whacker, Dave3457, DeadEyeArrow, Dirac66, DrEricYH, Dwmyers, Element16, Evercat, Fr1utbat, Herr blaschke, ILike2BeAnonymous, Icairns, Jammedshut, Jeodesic, Jiftzger, Johann Wolfgang, Karol Langner, Kinlee, Kjaergaard, Kkmurray, Loochsnuf, LostLucidity, Maartend8, Mark Oakley, Materialsscientist, Mboverload, Michael Hardy, Miguel Andrade, Mikaduki, Mikegreses, Mklewpatinond, Nakane, Nikai, Noosentaal, Obradovic Goran, PaddyM, Paolo.dL, PierreAbbat, RASnyder, Steve Quinn, Synchronism, The wub, Thorwald, Tldcollins, V8rik, WillowW, Zen Mind, 82 anonymous edits

Vibrational circular dichroism *Source:* http://en.wikipedia.org/w/index.php?oldid=354811869 *Contributors:* Aktsu, Auntof6, Bci2, Buurma, Dave3457, Indurand, LilHelpa, Petulda, R'nB, Slaweaks, Wnt, 3 anonymous edits

Optical rotatory dispersion *Source:* http://en.wikipedia.org/w/index.php?oldid=331863840 *Contributors:* BobbyBoulders, Chutznik, Dirac66, GTBacchus, Karol Langner, SemperBlotto, Sreffler, V8rik, YellowMonkey, 6 anonymous edits

Raman spectroscopy *Source:* http://en.wikipedia.org/w/index.php?oldid=352085509 *Contributors:* AJim, ARBradley4015, Afrine, Akv, Andrewavalon, Annabel, ArepoEn, Asfarer, Birdbrainscan, Blind cyclist, Brat32, Bullraker, Cdegallo, Charles Matthews, Christopherlin, Cybrol, D.Wardle, D3 TECHNOLOGIES, David Eppstein, Dazzaling69, Dch312, Dfbaum, Editore99, Fang Aili, Ferini, GT, Gabi bart, Gaius Cornelius, Galoubet, Gene Nygaard, Gene s, Gentgeen, Gerkleplex, GermanX, Gioto, Gunnar Larsson, Hankwang, Jaeger5432, Jaganath,

Jameslh, Janke, Jll, Jmameren, Jofox, Jonnyapple, Judenicholson, Keramamide, Kkmurray, Kraftlos, Kwamikagami, Latch.r, LordDamorcro, Loreshadow, LostLucidity, Lotje, MARKELLOS, MICKYGAL007, Magicalsauumy, Mahendra Kulkarni, Manulinho72, MarcoTolo, Martin Hedegaard, Measly Swan, Merope, Michbich, Mill haru, Minored, Mippi283, Moxfyre, Mythealias, Nihonjoe, Nikai, Nmnogueira, Paul August, Paul venter, Pavlina2.0, Pcarbonn, Pericles899, Petergans, Piano non troppo, Pixeltoo, Quantumobserver, RTC, Ravi khanna, Redleaf, Rich Farmbrough, Rob Hooff, Ronninget, Rosseth, Rulie123, Shashang, Shreevatsa, Smalljim, Srosie68, TDogg310, Tantalate, Tha Stunna, The number c, The wub, Thue, Tillwe, Tmb4bd, Tomatoman, Tomgally, Tzontonel, Uther Dhoul, Will4235, Wilson003, Yasuakinaito, Zylorian, 159 anonymous edits

Coherent anti-Stokes Raman spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=340339638> *Contributors:* Alash, Delldot, Eptoma, Feezo, Gwernol, HLOfferhaus, Interiot, Kkmurray, Lightmouse, MWS, Maartend8, Nerdeeksblonde, Nick Y., Onhoitsjamie, PigFlu Oink, Rich Farmbrough, ShakingSpirit, Signalhead, The wub, Thehelpfulone, V8rik, Vladisinger, 29 anonymous edits

Raman Microscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=308343741> *Contributors:* Gabi bart, Xezbeth

Imaging spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=346629165> *Contributors:* Andyphil, Bci2, Curaci, Donarreiskoffer, HYPN2457, Idg101, Ivan Shmakov, JFBolton, Jaeger5432, Jim.henderson, Kkmurray, Larryloz, Ligulem, NathanHagen, Nbarth, Omegatron, Pearle, Pjvpjv, Schae001, Steve Quinn, ThaddeusB, Wadje, Zowie, 22 anonymous edits

Chemical imaging *Source:* <http://en.wikipedia.org/w/index.php?oldid=352448633> *Contributors:* Alansohn, Andyphil, AngelOfSadness, Annabel, Banus, Batykefer, Bci2, BierHerr, Chris the speller, Closedmouth, D6, Davewild, Editore99, Fgnievinski, Gabi bart, GeeJo, HYPN2457, Iridescent, JIP, Jim.henderson, Kkmurray, Larryloz, Mdd, Mkansiz, Natalie Erin, Skysmith, Stone, Tassedethe, Ultraexactzz, Wilson003, 40 anonymous edits

Spin polarization *Source:* <http://en.wikipedia.org/w/index.php?oldid=337393318> *Contributors:* Arnero, Ctchiang, Kusma, Michael Hardy, Mindmatrix, RDR, SCEhardt, Shaddack, Tayga, V8rik, Zawersh, 4 anonymous edits

Polarized Neutron Spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=315969727> *Contributors:* Cardamon, Chrisainthere, EBlackburn, Eubene, Joachim Wuttke, Neelix, Tpkonen, 3 anonymous edits

Polarized Muon Spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=353396210> *Contributors:* BWDuncan, Dalibor Bosits, DanMS, Galaxiaad, Gene Nygaard, Grj23, JHBrewer, Jcwfi, Kite0419, Kusma, Materialscientist, Pionade, RDR, Richtea2007, Sergio.ballestrero, Slavomirkapusta, Snafu450, Strait, Thinking of England, WikHead, 20 anonymous edits

Time-resolved spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=353725604> *Contributors:* Azo bob, Cenarium, Christopherlin, GregorB, Hankwang, Man It's So Loud In Here, SJP, THEN WHO WAS PHONE?, 12 anonymous edits

Terahertz spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=323341545> *Contributors:* Frankhindle, Johnlp, Kkmurray, Rich Farmbrough, Zroutik, 1 anonymous edits

Applied spectroscopy *Source:* <http://en.wikipedia.org/w/index.php?oldid=352902543> *Contributors:* Bambika, Draganakusiclazic, Egpetersen, Graeme Bartlett, Headbomb, Itub, Jaeger5432, Kkmurray, Millosh, Peterlewis, Pro crast in a tor, Sbialkow, Woohookitty, 6 anonymous edits

Amino acids *Source:* <http://en.wikipedia.org/w/index.php?oldid=16099379> *Contributors:* 134.95.200.xxx, ABCD, Aardtek, Aboveleft, Acroterion, Adashiel, AdjustShift, Advanet, AgentCDE, Ahoerstemeier, Aitias, Alansohn, Amdk8800, Amenray, Anclation, AndonicO, Andre Engels, Andres, Animum, Antandrus, Antiuser, Antorjal, Apodtele, Arakunem, Arcadian, Archer7, Ask123, Astronomicalunit, Atemperman, AtheWeatherman, Ausinha, AzaToth, Azzamsyr, Baa, Beetstra, Ben-Zin, Bender235, Bendzh, Benjah-bmm27, Bennnh, Bensaccount, Berkzafer, BernhardH, Betterusername, Bettingr, BIT, Billfred, Bisected8, Blacksmith tb, Blanchardb, Bobo192, Bomac, Bongwarrior, Borb, BorisTM, Bornhj, Breno, BruceDLimber, Bryan Derksen, Cacycle, Can't sleep, clown will eat me, Captain panda, Cbeau, Cenarium, Ceyockey, Chem-awb, Chino, Chirpy7, Chris Cappaia, Cless Alvein, Clicketyclack, Cliff smith, ClockworkSoul, Colonies Chris, Cometstyes, Computerjoe, Conversion script, Cooleshunk, Coppertwig, Coredesat, CorpX, Cpeditorial, Cpl Syx, Crazycomputers, Crookshank2227, Crum375, Cubathy, Cwenger, Cyclonenim, DMacks, DVD R W, DWIII, Dabigkid, Dabomb87, Dan Wylie-Sears, Dancojocari, Daniel5127, Darth Panda, Deli nk, Delta G, Dfrg.msc, Diberri, Discospinster, Dmb000006, Dogposter, Donarreiskoffer, Donjavi, Dprust, Dreg743, Drini, Dungodung, Duplode, Dwmyers, Dysepsion, EconoPhysicist, Ectropist, Edgar181, Eequor, Eggwadi, El C, Elassin, Elb2000, Elbodno, Eleassar777, Emw, Enchanter, Encyclopedia77, End3rtheThird, Eprh123, Eribro, Erickulcyk, EscapingLife, Evercat, Everyking, Excirial, Faradayplank, Fedir, FeldBum, FelixP, Fmandog85, FreeKresge, Freddie, Fritzpoll, Fubar Obfusco, Fullstop, G. Campbell, G3pro, Galizur, Gentgeen, Geoffreynham, Giftlite, Gilliam, Glenn, God Emperor, Gurch, Gzkn, Habj, Hadal, Hajatvrc, Harris7, Harryboyles, HoergerJ, Hsn mhd, Hybrid04, II MusLiM HyBRiD II, IRP, IW.HG, Ihope127, Immunize, IrisKawling, It Is Me Here, Itub, Ixfid64, J04n, JForget, JNW, JWBE, James McNally, Jamil17, Javert, Jaw959, JediBookworm, Jeedpay, Jeffrey Mall, JeramieHicks, JeremyA, JesseGarrett, Jfdwolff, JimJ, Jock Boy, Joe Beaudoin Jr., John, Jomunro, Josh Cherry, Kaarel, Kammler1, Kandar, Karl Dickman, Kartano, Keenan Pepper, Keilana, Kengo1040, Kevinallencrittenden, Kgasso, Khajidha, Khooly59, Kijitah, Kjkolb, Knivusa, Knutux, Kostisl, Kpjas, Kubra, Kuru, L-Tyrosine, La goutte de pluie, Leeyc0, Leszek Jańczuk, Lexor, Lights, LikeHolyWater, LilHelpa, Ling.Nut, Llull, Looxix, LosDorks, Lovey73, LuckyWizard, Lumos3, Miss1ontomars2k4, MD87, Maartend8, MacGyverMagic, Magnus Manske, Malljaja, Mani1, Marj Tiefert, Master gopher, Matt Fitzpatrick, Matteo2000, Matthew Yeager, Mattisse, Maurice Carbonaro, Mav, Maxamegalon2000, Maximaximax, Mayalld, McGeddon, Memset, Mike2vil, Molybdenumblue, Monurkar, Moryoav, MrOllie, Mushin, Myanw, N5iln, NHRHS2010, Naive cynic, Nandesuka, Narayanese, Natalie Erin, Naturespace, Nevin Janzen, Nihilres, Nikki311, Nilfanion, Nlu, Nomad2u001, Nono64, Nucleophilic, Nuklear, Oasissbob, Obradovic Goran, Olivier, Omegatron, Onco p53, Opabinia regalis, Openlander, Orangutan, Otisjimmy1, Otiavee, Oysteroadfish, Pansi gupta, Patstuart, Pauli133, Peak, PeterSymonds, Philip Trueman, Physicm82, PierreAbbat, Possum, Prilla, Prodego, Proofreader77, Qef, Qqx, Rada, Radon210, Res2216firestar, Rewster, RexNL, Rich Farmbrough, Richard Arthur Norton (1958-), Richard001, Rifleman 82, Rittard, Rleel185, Rmosler2100, Roadrunner, Roadsoap, Robna, Rory096, RoyBoy, Sam Hoecevar, SchnitzelMannGreek, Seans Potato Business, Selket, Serinde, Shaddack, ShadowBlade1000, Sheehan, Shekharsuman, Sikkem, SimonP, Sionus, SirGrant, Skoddet, Slgoeki, Slow Riot, Slp1, Smithbrenon, Smokefoot, Snowmanradio, Soren.harward, SpLoT, SpaceFlight89, Squidionius, SriMesh, Stephentaylor, Stewartadcock, Stillnotelf, Stismail, Stone, Stormhawk, Superarthur, Switchsonic, Sydney naturopath, TBadger, Tarawneh, Tarotcards, Teddks, Teles, Thaseint, Thehelpfulone, Tiggerjay, TimVickers, Tjod, Tobias Hoevekamp, Tonderai, Tony1, TonyTheTiger, Tounos05, Toyz, Transhumanist, Tremello22, Trevor Wrenblom, Tristanb, Trlbizr, Truthflux, Ttomy21, Tuckerekcult, Turnstep, Tyler, Ultimate sickness, Umleashuh, Unukonoro, User A1, Ustye, Utcursch, Uttam Pal, Versageek, Vina, ViolentRage, WVhybrid, White.matt.09, Whittlewiki, Wiki007wiki, William Avery, WillowW, Wimt, Winchelsea, Wizzy, Wolfmankurd, WriterHound, Wrstroud, Wtmitchell, Ww, Wwallacee, Xinit, XoiLov3Yououx055, Xook1kai Choa6aur, Yamamoto Ichiro, Yoasif, Yobmod, Yyy, Zaphraud, Zhang He, Zoicon5, -K, 836 anonymous edits

Proteins *Source:* <http://en.wikipedia.org/w/index.php?oldid=232130474> *Contributors:* 0, 162.129.26.xxx, 168...., 200itlove, 8472, A K AnkushKumar, A-giau, AA, ABF, Aarktica, Abce2, Achilles.g, Acroterion, ActivExpression, AdamRetchless, Adashiel, Adenosine, AdjustShift, Agent Smith (The Matrix), Ageo020, Agesworth, Agilemolecule, Ahoerstemeier, Aitias, Alansohn, Ale jrb, Alexjohnc3, Alison, Alphachimp, Altruistic Egotist, Alvestrand, Amozt, Ams80, AnOddName, Anclation, AndonicO, Andre Engels, Andrewpkmz, Andries, Angr, Annabel, Antandrus, Anthony, Aragorn2, Arakunem, Arcadian, Arkuat, Ascidian, Ashdog137, Asher196, Aspenandgwen, AssistantX, Astrojan, AuburnPilot, AutoFire, AxelBoldt, Ayleuss, Aznerd123, BW52, BadSeed, Banes, Banks, Banus, Barticus88, Beano, Beetstra, Belovedfreak, Ben-Zin, Ben.c.roberts, Benbest, Benbreadd, Bender235, Bendzh, Bensaccount, Bernardbd, Bulsepaga, Biblee, BigCow, Bihero, Bioinformin, Biophys, Bloodshedder, Bobo192, Bogdan, Boghog2, Bomac, Bongwarrior, Bookandcoffee, BorisTM, Bornhj, Brian the Editor, Brigand dog, Brisevegas, Bryan Derksen, Btharper1221, Bubbachuck, Burzmali, Clreland, CWY2190, Cacycle, Caltas, CambridgeBayWeather, Can't sleep, clown will eat me, CanadianCaesar, Capitana, Capricorn42, CardinalDan, Carl Caputo, Carl jr., Cassandra 73, Cburnett, Cenarium, Centrx, Champ0815, Chasingsol, Chenmengen, ChicXulub, Chino, Christoph.gille, Christopher Parham, Chuck Carroll, Citicac, DanCC, Clarince63, Clicketyclack, ClockworkSoul, Closedmouth, Cmdrjharmonie, Commanderigg, Conversion script, Coookeee crisp, Coppertwig, Corti, Cpeditorial, Cpunwhiz11, Cryptic, Crystallina, Cst17, CupOfRoses, Curehd, Curps, Cyclopia, Cyde, DMacks, DSachan, DVD R W, Dabomb87, Dan Gagnon, Danizdeman, Dbarker348, Dbfirs, De728631, DeadEyeArrow, Debresser, Deli nk, Delldot, Delta G, DennisDaniels, DerHexer, Derekwriter, Derezo, Dgies, Dina, Discospinster, Dlohcierekim, Dmb000006, DocSigma, Doctor Knoooow, Doctorcherokee, Donarreiskoffer, Douglasmattaylor, Drdozer, Dread Specter, Drini, Drphilharmonic, Dthzip, Ducttape17, Duckync2, Dureo, Dwmyers, Dycedarg, Dysepsion, D6Rahier, ERcheck, ESkog, EconoPhysicist, Edgar181, Editore99, Eequor, Eganio, El C, Elagatis, Electrified mocha chinchilla, Emw, Enviroboy, Eprh123, Eric Forste, Eric Martz, Eric-Wester, Erik Zachte, Espoo, Essayjay, Euchiasmus, Everyking, Exir Kamalabadi, Fenteany, Fieldday-sunday, Fittedneseducation, Fjool, Flavonoid, FocalPoint, Frank 212121, Fratrep, Fresheneesz, Frictionary, Fruge, FuelWagon, Fvasconcellos, G. Campbell, G3pro, GBoran, Gail, Gaius Cornelius, Galilee12, Galoubet, Gangsta4lif, Gene Nygaard, GeneralIroh, Giftlite, Gilgamesh he, Gilliam, Gimmetrow, Glane23, Glen, Glenn, Gogo Dodo, GraemeL, GrahamColm, Grimey109, Gurch, Hadal, Haham hanuka, Hangjian, Hasek is the best, Hdt83, Hedgehog33, Hemmingsen, Henning Blatt, HernanQB, HexaChord, HistoryStudent113, Holderca1, Holeung, Husond, I b pip, II MusLiM HyBRiD II, IWhisky, Icek, Ihope127, Ike9898, Indosauros, Insanephantom, Interiot, Iridescent, Irishguy, Isopropyl, Itln boy, Ivan Akira, Ivo, Iwearsox21, Ixfid64, J. Finkelstein, J.delanoy, JForget, JSpudeman, JWSchmidt, Jackol, Jake11, James086, James42.5, Jamesontai, Janejellyroll, Jasonzhuocn, Javanbakti, Jbeans, Jecar, JeremyA, Jerryseinfeld, JesseOjala, Jhs1681, Jipan, Jj137, Jlaudiow713, Jls043, Jman5, JoanneB, Joe Jarvis, Joeywallace9, Johan Elisson, John254, Johnchiu, Jojit fb, JonReader, Jonathan Grynspan, Jordell 000, Jose77, Josh Cherry, Jpers36, Jstraining, Jkiefre, Juliancolton, Jusjih, Kalyandachkravarthy, Karen Johnson, Karlhahn, Karol Langner, Katherine Folsom, Kbh3rd, Keilana, Keith Lehwald, Kenschaft, Khaledelmansoury, Khooly59, King Lopez, Kingialis3, Kingpin13, Kkmurray, Knightdude99, Knowledge Seeker, Knutux, Korg, Kubigula, Kukini, Kyawtun, KyraVixen, Kznf, LfRaone, Lafw, Lakers, Landon1980, Lantonov, Lascorz, Lawilkin, Laxman12, Leafyplant, LeaveSleeves, LedgendGamer, Leebo, Lemchesvej, Lexor, Lindmere, Lir, Lordfeepnuss, Luna Santin, Lunschscale, Lupo, M412k, MBisanz, MC10, MECU, MEJ119, MER-C, Mac, Magioladitis, Magnus Manske, Mai-tai-guy, Maianwilliams, Malatesta, Malcolm Farmer, Manojdhwade, Marek69, Marj Tiefert, MarkSutton, Martin451, Masterofsuspense3, Matticus78, Matturm, Mattwolf7, Mav, Mbecker, Mccreedy, Megalodon99, Megan1967, Melsaran, Mentisock, Mephistophelian, Merovingian, Meruem, Mfickes, Michael Hardy, Michaels10, Mifter, MightyWarrior, Mike2vil, Mike713, Mikewall, Mikez, Miller17CU94, Minimac's Clone, Minorfixer, Miss Madeline, Missellyah, Misza13, Mkdw, Mkubica, Mkv22, Mlsquirrel, Montchav, Moonriddengirl, Moshe Constantine Hassan AI-Silverburg, Mr. Lefty, Mr.Bip, Mr.Z-man, Msabic, Muad, Munita Prasad, Mxn, My very best wishes, Mygerardromance, NClement, NERIC-Security, Nadyes, Nakon, Narayanese, Nat11, Natyredude858, Naturespace, NavlinWiki, Ndkartik, Neil.steiner, Neile, Netesq, Netoholic, Neurolysis, NewtN, Nibuod, Niceguyedc, Nick Garvey, Nick Van Ni, Nitin77, NuclearWarfare, Nugneant, Nuttycocoanut, ONUnicorn, Oda Mari, OldakQuill, OllyG, Omicronperse8I, Omnipedian, Onco p53, Onorem, Opabinia regalis, Ossmann, Outrigger, Oxymoron83, P Carn, P99am, PDH, PandaDB, Paranoia, Paranthaman, Password,

Pdcook, Peak, Pekinensis, Perfecto, Persian Poet Gal, Peruvianllama, Pkg, Pharaoh of the Wizards, Phgao, Philip Trueman, Piano non troppo, Picaroon, Pierpunk, PierreAbbat, Pinethicket, Pipe34, Plankwalking, Player 03, Poindexter Propellerhead, Polluks, Pookythegreat, Possum, Ppntori, Pravinhiwale, Priscus, Public Menace, Pyrospirit, Quadzilla99, Qxz, R. S. Shaw, RTucker, Rada, Rama's Arrow, Raven4x4x, Redvers, Reo On, RexNL, Rich Farmbrough, Richard Arthur Norton (1958-), Richard Taylor, Richfife, Rickert, Rmky87, RoadieRich, Roadsoap, Robert Foley, Rockvee, Rodhullandemu, Roland2, Romanm, Ron Ritzman, Ronjhones, Rrburke, Rror, RxS, RyanGerbil10, SJP, Sam1001, Samsara, Samwebstah, Sardanaphalus, Sasata, Satyrium, Schwjn, Scohous, ScotchMB, Seaniemasters, Seba5618, Sentausa, Setters, Sepsaros, ShakingSpirit, Shanel, Sharpvisuals, Shella, Shinmawa, Shizhao, Shoessas, Shura58, Sid 3050, Siim, Sijo Ripa, Sir Leks, Sirsanjuro, Sjakalle, Slashme, Snowmanradio, Snoyes, Sonic h, Soumyasch, Special-T, Spiff, Spritegenie, SpuriousQ, Squidonius, Srice13, STAKAR Karnak, StaticGull, Staxringold, Stelligent, Stephenb, Stewartadcock, StuartF, Stubbyhead, Studerby, Subsolar, Suffusion of Yellow, SunDragon34, Sunray, Super-Magician, Superbeecat, Susvolans, Swarm, Sysy, THF, Tad Lincoln, Tameeria, Taoster, Tarquin, Tatterfly, Taw, Tawker, Taxman, Tellyaddict, Tempodivalse, Tennispro427, TestPilot, Tetracube, That Guy, From That Show!, The Hokkaido Crow, The Rambling Man, The Thing That Should Not Be, Think outside the box, ThinkerThoughts, Thinking Stone, This doesnt help at all, Thishumorcake, Tide rolls, TigerShark, Tim1357, Tim2027uk, TimVickers, Timdownie, Timir2, Titoxd, Tizio, Toby Bartels, Tony1, Tpkonen, Tprentice, Trigger hippie77, Triwbe, Tsemii, Ttiotsw, Turnstep, TutterMouse, Twilight Realm, Tyciol, Tylerhammond2, Uirauna, Uncle Dick, User A1, Usmanmyname, VK3, VasilievVV, Vector Potential, Versus22, Voice of All, WLU, Wardface, WarthogDemon, Wayne Miller, Wayward, Whkoh, Wiki alf, Wikimercenary, WikipedianMarlith, WillowW, Wimt, Winchelsea, WinterSpw, Wisco, Wk muriithi, Wknight94, Wolfkeeper, Woohookitty, Work permit, Wuhwuzdat, Wwood, Xiaosflux, Xavierschmit, Xelixed, Yamamoto Ichiro, Yanggers, Yidisheryid, Yurilinda, Yyy, Zafiroblue05, Zarius, Zenzhu, Zephyris, Zoescow, ZooFari, Zsinj, Zzuuzz, 1612 anonymous edits

Protein structure *Source:* <http://en.wikipedia.org/w/index.php?oldid=354461300> *Contributors:* AC+79 3888, Abaughv, Ahoerstemeier, Akjohnson, Alan Au, Alansohn, Albal, Alchemist Jack, Anonymi, Antony-22, Apparition11, Arcadian, AxelBoldt, Banus, Beno1000, Berkly, Biophys, Bluskyfing, Bobo192, Can't sleep, clown will eat me, CardinalDan, Choij, Chris 73, Christopherlin, Chymera, Clicketyclack, Cmdrjameson, Cryptic C62, DVD R W, Dcrjsr, Debresser, DerHexer, Dmb000006, Element16, Elmer Clark, Enviroboy, Etrxge, Euchiasmus, Foolip, Friginator, Gaius Cornelius, Gene Nygaard, Ghakko, Gilliam, Graeme Bartlett, Grafened, Gurchzilla, Hbent, Ike9898, Jdelanoy, JaGa, Jan1nad, Jcwf, Jordell 000, K3rb, Kanonkas, Kjaergaard, Kochipoi, Konstantin, La Pianista, Law, Lexor, Loodog, Lynnbridgebook, Malenkyilzards, Mandolinface, Math-gamhainn, Michael A. White, Miguel Andrade, Mikel, Moleculeoflife, Mushin, NeoJustin, Nina Gerlach, Nitya Dharma, Omicronpersei8, Opabinia regalis, P99am, Password, Pdcook, Phgao, Pinethicket, Pravinhiwale, Pro crast in a tor, Q31245, ROBE0191, Rasmusw, Recury, Retama, Rkiriarn, RyanGerbil10, Saganatsu, SeventyThree, Shadowjams, Shrimp wong, Sir Vicious, Sluzzelin, Smoe, Speedyboy, Splette, Stefano Garibaldi, Sverdrup, Team6and7, Theyeti, Thorwald, Tim Ross, TimVickers, Tommymac10, Tralala, Trusilver, Vijaykumarutkam, Vriend, Webridge, Wheedhee, Whiner01, WillowW, Yamamoto Ichiro, 255 anonymous edits

Protein folding *Source:* <http://en.wikipedia.org/w/index.php?oldid=354827918> *Contributors:* 168...., 5beta5, Adriferr, Agilemolecule, Akane700, Andraaide, Antony-22, Arcadian, Artoannila, Banus, Baticus88, Bci2, Bendzh, Bfinn, Bioinfo177, Biophys, Biophysik, Bkhouser, Blainster, Blooooo, BlurTento, Brianga, Bryan Derksen, Cacycle, Calvero JP, Calthagarvey, Cburnett, ChicXulub, Clicketyclack, Computor, Cyberman, Czhangrice, D. Recorder, DannyWilde, Dave3457, Davepnr, DennisDaniels, Dhatz, Donarreiskoffer, Dwmyers, ESKog, Eequor, Erencexor, Erwinser, Fawzin, Fuzheado, Gcrossan, Gowantervo, Grimlock, Harley peters, Herd of Swine, Hooplehead, Intangir, Ixf6d44, JJ TKOB, Jacobsman, Jammedshut, JeramieHicks, Katherine Folsom, Keyvin, Kierano, Kjaergaard, Konstantin, Kostmo, LarRan, Leptictidium, Lexor, Lfh, LiDaobing, Lir, Lostart, Lucaah, M stone, Macintosh10000, Madeleine Price Ball, Magnus Manske, Malcolm Farmer, Mark Renier, Michael Hardy, Miguel Andrade, Minghong, Mittinatten, MockAE, Movado73, Movalley, My very best wishes, Myscrnm, Netesq, Opabinia regalis, Otvaltak, P99am, Piotrus, Pro crast in a tor, Ptr123, RDBrown, Rebroad, Rich Farmbrough, Richyfp, Rjwilmsi, RoyBoy, RunninRiot, Samrat Mukhopadhyay, Sharkman, Shmedia, SirHaddock, Smelialichu, Snowmanradio, Splette, SpuriousQ, Stepa, Sunny17152, Tan, Terrace4, TestPilot, The Anewe, The Great Cook, Tijmz, TimVickers, Tomixdf, Tommstein, Toytoy, Trenchcoatjedi, Trusilver, V8rik, Waerloeg, Wavelength, WillowW, WorldDownInFire, Xnuala, Xy7, Yves-henri, Zoicon5, גוראיר, 168 anonymous edits

Protein dynamics *Source:* <http://en.wikipedia.org/w/index.php?oldid=290533592> *Contributors:* 168...., Ayzov, Altenmann, Anthony Appleyard, Bhadani, Biophys, Bulgaroctonus, CRGreathouse, Colonies Chris, Dejo, Dicky2206, Doopa, Element16, Flyguy649, Hoffmeier, Johnpseudo, Jvbishop, K.murphy, KestasUPSY, Loupeter, Mary Ann Summers, Mikael Häggström, MockAE, Nneonneo, Noalaignorancia, NuclearWarfare, P99am, Raffuzzo p, Rajah, Richard001, Sciencetalks, Sintaku, Someche, Substatique, Tameeria, Thorwald, Tibor Horvath, Timwi, Txnman307, Whiner01, Whosaking, Zashaw, 39 anonymous edits

Nucleic Acids *Source:* <http://en.wikipedia.org/w/index.php?oldid=20758770> *Contributors:* 168...., 2 of 8, A4, Abrech, Acelgoyobis, Adam Bishop, Addshore, Agathman, Alboyle, Alex43223, Alexius08, Anclation, Andre Engels, Angela, Anthere, Antony-22, Antzervos, Aoi, Arakunem, Araucaria, Arbitrarily0, Babyranks17, Barticus88, Beethoven's DNA, Beetstra, Bensaccount, Bobchicken9, Bobo192, Bradjamesbryan, CDN99, CL, CLW, Cacycle, Can't sleep, clown will eat me, Careless h, Ceyockey, Chanoyu, ChaosR, ChrisCork, ChrisHodgesUK, Ciphers, Conversion script, Cpeditorial, Cyn, Daniel5127, DarkFalls, DerHexer, Desireader, Discospinner, Dng267, Dposse, Dungodung, Dysprosia, EconoPhysicist, Ed Poor, Ellmist, Emw, Forluvoft, Frankenpuppy, Franz Bryan, FreplySpang, G3pro, Gaia Octavia Agrippa, GeeJo, Gentgeen, Giftlite, Gilliam, Gogo Dodo, GraemeL, Graft, Grunty Thraveswain, Gurch, Gwernol, Hede2000, Hellow yo, Hellowhy, Hephaestos, Hu, Hydrogen Iodide, I dream of horses, It Is Me Here, J Di, J.delanoy, JForget, James086, Jhannah, Joanjoc, John R Murray, JonMoulton, Josh Cherry, Josh Grosse, K3f3rn, Kablammoo, Kandar, Kbdank71, Keegan, Kierano, KimvdLinde, Kkmurray, Kpjas, Kyle 290, LAX, La goutte de pluie, Laurascudder, Leonard Vertigheal, Lexor, Liamdaly620, Lir, Looie496, Luna Santin, LyXX, M80forYOU, Magog the Ogre 2, Marudubshinki, Mav, Maxxicum, Mendaliv, Miaow Miaow, Mikegrant, Misza13, Mlouns, Mygerardromance, Narayanese, Nasril, Ngourlie, Nk, Npss, NuclearWarfare, Omicronpersei8, Onco p53, OrangeDog, Oxymoron83, P99am, Pakaran, Pgan002, Pharaoh of the Wizards, Philip Trueman, Piano non troppo, Pobregatica, Possum, Ppbgardine, Ppntori, Preston.hussey, Pschemp, RJaguar3, Recej, Res2216firestar, Rifleman 82, SJP, Scharks, Schmutz MDPI, Sciurinae, Seba5618, Sgsparos, ShellCoder, Silsor, Simeon H, Sintaku, SixPurpleFish, Smack, Smaug123, Snoyes, Squidonius, Stephen5406, Stewartadcock, Str1977, Strathallen, Tad Lincoln, Talon Artaine, TheTito, Thehelpfulone, Thingie, Tide rolls, Tim1357, TimVickers, TimonyCrickets, Trevor MacInnis, Tuganax, U.S.A.U.S.A., Ugen64, Uncle Dick, User A1, VIOLENTRULER, Vary, Versus22, Violetriga, W3bj3d1, Waggors, Wikipedia addict101, Wimt, Wisdom89, Wysprgr2005, Xdenzien, Yachtsman1, Yk Yk Yk, Ziusudra, 488 anonymoud edits

DNA *Source:* <http://en.wikipedia.org/w/index.php?oldid=355095288> *Contributors:* (, jarbarf), -Majestic-, 168...., 168...., 169, 17Drew, 3dsceience, 4u1e, 62.253.64.xxx, 7434be, 84user, A D 13, A bit iffy, A-giau, Aaaxlp, Aatomic1, Academic Challenger, Acer, Adam Bishop, Adambiswanger1, Adamstevenson, Adashiel, Adenosine, Adrian.benko, Ahoerstemeier, Aitias, Aj123456, Alai, Alan Au, Aldaron, Aldie, Alego092, Alexanderemas, Alkivar, Alphachimp, Alzhaid, Amboo85, Anarchy on DNA, Ancheta Wis, AndonicO, Andre Engels, Andrew wilson, Andreww, Andrij Kurtsky, Andjycp, Anita1988, Anomalocaris, Antandrus, Ante Aikio, Anthere, Anthony, Anthony Appleyard, Antilient, Antony-22, Aquaplus, Aquilla34, ArrezZeynili, Arcadian, Ardyn, ArielGold, Armored Ear, Artichoker, Asbestos, Astrowob, Atlant, Aude, Autonova, Avala, AxelBoldt, AyS288, AzaToth, B, BD2412, BMF81, Banus, BaronLarf, Bbatsell, Bci2, Bcorr, Ben Webber, Ben-Zin, BenBildstein, Bender235, Benjah-bmm27, Bensaccount, Bernie Sanders' DNA, Bevo, Bhadani, Bhar100101, BiH, Bijee, BikA06, Bill Nelson's DNA, Billmccnt189, Biolinker, Biriwilk, Bjwebb, Bkell, Blastwizard, Bloger, Blondtrallite, Bmtbomb, Bobblevik, Bobo192, Bongwarrior, Borislue, Bornjh, Brian0918, Brighterorange, Briland, Brim, Brockett, Bryan, Bryan Derksen, CWY2190, Cacycle, Caerwine, Cainer91, Cal 1234, Calaschysm, Can't sleep, clown will eat me, Canadaduane, Carbon-16, Carcharoth, Carlo.milanesi, Carlwev, Casliber, Calthagarvey, CatherineMunro, CattleGirl, Causa sui, Cburnett, Cerberus lord, Chanora, Chanting Fox, Chaojoker, Charm, Chill Pill Bill, Chino, Chodges, Chris 73, Chris84, Chuck Grassley's DNA, Chuck02, Clivedelmonte, ClockworkSoul, CloudNine, Collins.mc, Colorajo, CommonsDelinker, Conversion script, Cool3, Coolawesome, Coredsat, Cornacchia123, Cosmotron, Cradlelover123, Crazycomputers, Crowstar, Crusadeonilliteracy, CryptoDerk, Crzussian, Cubsrazy29, CupOBeans, Curedh, Curps, Cyan, Cyclonenim, Cyrius, D6, DIREKTOR, DJAX, DJRafe, DNA EDIT WAR, DNA is shy, DVD R W, Daniel Olsen, Daniel987600, Danielkueh, Danny, Danny B-), Danski14, Darklilack, Darth Panda, Davegrupp, David D., David Eppstein, Davidbspalding, Daycd, Db099221, Dbabbitt, Dcoetzee, DDeceShooter, DeadEyeArrow, Delltdo, Delta G, Deltabeignet, DevastatorIIC, DiBerri, Dicklyon, Digger3000, Digitalme, Dina, Djm1279, Dlohcierekim's sock, Dmn, DocWatson42, Docjames, Doctor Faust, Docu, Dogpostor, DonSiano, Donarreiskoffer, Dr d12, Dr.Kerr, Drini, Dudewheresmywallet, Dullhunk, Duncan.franz, Dungodung, Dysmorodrepanis, E. Wayne, ERcheck, ESKog, Echo park00, Echuck215, Eddycrai, Editing DNA, Edwy, Efbfweborg, Egil, ElTyrant, Elb2000, Eleassar777, EliasAlucard, ElinorD, Ellmist, Eloquence, Emoticon, Emw, Ephemerium, Epingchris, Erik Zachte, Escape Artist Swyer, Esumrin, Etanol, Etrrig, EurekaLott, Everyking, Evil Monkey, Ewaser, Excavator, FOTEMEHL, Fabhcuín, Factual, Fagstein, Fastfission, Fconaway, Fcrick, Fernando S. Aldado, Ffirehorse, Figma, Figure, Firetrap9254, Fishingpal99, Flavaflav1005, Florentino floró, Fnielsen, Forluvoft, Freakofnurture, FreplySpang, Friendly Neighbour, Frostyservant, Fruge, Fs, Fvasconcellos, G3pro, GATHrawn22, GHe, GODhack, Gaara san, Galoubet, Gary King, Gatorcpk, Gazibara, GeeJo, Gene Nygaard, GeoMor, Giftlite, Gilisia, Gilliam, Gimmetrow, Gjuggler, Glen Hunt's DNA, Glenn, Gmaxwell, GoEThe, Goatasaur, Gogo Dodo, Golaazofotahabadi, GordonWatts, Gracenotes, Graeme Bartlett, GraemeL., Grafikm fr, Graft, Graham87, GrahamColm, Grandegrandegrande, GregorB, Grover cleveland, Gurko, Gustav von Humpelschmumpel, Gutza, Gwsrinme, Hadal, Hagerman, Hairchrn, Hairwheel, Hammersoft, Hannes Röst, Harianto, Haylejohnson21, Heathhunnicutt, Hephaestos, Heron, Heyheyhack, Hockey21dude, Horatio, Hu, Hughdbrown, Hurricanehink, Hut 8.5, Hvn0413, I hate DNA, Iapetus, Icairns, Iai Kr., Imppamiizgraa, InShaneeen, Inge-I.yubov, Isilanes, Isis07, Itub, Ixf6d44, Izehar, JHMM13, JWSchmidt, JWSurf, Jacek Kendysz, Jackrm, JamesMLane, JamesMt1984, Janejellyroll, Javert, Jaxl, Jeka911, Jer ome, JeremyA, Jerzy, Jetsetpainter, Jh51681, Jiddisch, Jimriz, Jimwong, Jlh29, Jls043, Jmcc150, Jo9100, JoanneB, Joconnol, Joeywallace9, Johansv, JohnArmagh, Johtext, Johnniucj, Jojit fh, JonMoulton, Jonrunles, Jonverve, Jorvik, JoshuaZ, Josq, Jossi, Jsteh, Julian Diamond, Jumbo Snails, Junes, Jwrosenzweig, Kblfin, Kapow, Karmman, Kazkazkazkasako, Kbh3rd, Keegan, Keepwee, Keilana, Kelly Martin, Kemyou, Kendrick7, Kerry077, Kevin B12, Kevmitch, Kghose, Kholdstare99, Kierano, KimvdLinde, King of Hearts, KingTT, Kingturtle, Kitch, Knowledge Seeker, KnowledgeOfSelf, Koavf, KrakatoaKatie, Kums, Kungfuadam, Kuru, Kwamikagami, Kwekubo, KyNephi, LA2, LFaraoe, La goutte de pluie, Lascorz, Latka, Lavateraguy, Lee Daniel Crocker, Lemchesvej, Leptictidium, Lerdsuwa, Leuko, Lexor, Lhenslee, Lia Todua, LightFlare, Lightmouse, Lightspeedchick, Ligulem, Lincher, Lion Wilson, Lir, Llongland, Llull, Lockesdonkey, Logical2u, Loginbuddy, Looxix, Loren36, Loris, Luigi30, Luk, Lumos3, Luna Santin, Luuva, MER-C, MKoltnow, MONGO, Mac, Madeleine Price Ball, Madhero88, Magadan, Magnus Manske, Majorly, Malcolm rowe, Malo, Mandyj61596, Mantissa128, Marcus.aerlous, Marj Tiefert, Martin.Budden, MarvPaule, Master dingley, Mattbr, Mattbrundage, Mattjbylte, Maurice Carbonaro, Mav, Max Baucus' DNA, Max Naylor, McDogm, Medessec, Medos2, Melaen, Melchior, Mentalmaniac07, Mgiganteus1, Mgoohey, Mhking, Michael Devore, MichaelHa, Michaelas10, Michigan user, MidgleyDJ, Midnightblueowl, Midoriko, Mika293, Mike Rosoff, Mikker, Mikko Paananen, Mintman16, MisfitToys, Misza13, Mithent, Mjpieters, Mleefs7, Moink, Moorice, Mortene, Mr Bungle, Mr Meow Meow, Mr Stephen, MrErku, Mstislavl, Mstroeck, Mulad, Munita Prasad, Muro de Aguas, Mwanner, Mxn, Nakon, Narayanese, Natalie Erin, Natarajanganesan, Nate1028, NatureA16, Nauseam, Nbauman, Neckro, Netkinetic, Netoholic, Neutrality, Never give in, NewEnglandYankee, Nighthawk380, NighthawkJ, Nhiltres, Nirajim, Nishkid64, Nitecrawler, Nintarekcap, No Guru, NoldeaNick, NochnoiDozor, Nohat, Northfox, NorwegianBlue, Nthornberry, Nunh-huh, OBloodyHell, OODDDD, Obli, Oblivious, Ocolon, Ojl, Omicronpersei8, Onco p53, Opabinia regalis, Opelio, Orrin Hatch's DNA, Orthologist, Ortolan88, Ouishoebean, Outgrgr, OwenX, P99am, PDH, PFHLai, PaePae, Pakaran, Pascal666, Patrick0Moran, Patrick2480, Patstuart, Paul Foxworthy, Paul venter, Paulinho28, Pcb21, Pde, Peak, Pedro, Persian Poet Gal, Peter Isotalo, Peter K., Peter Winnberg, Pgan002, Philip Trueman, PhilipO, Phoenix Hacker, Pierceno, PierreAbbat, Pigman, Pigmietheclub, Pilotguy, Kirlin, Poor Yorick, Portuguese6927,

Potatoswatter, Preston47, Priscilla 95925, Pristontale111, Pro crast in a tor, Prodego, Psora, PsyMar, Psymier, Pumpkingrower05, Pyrospirit, Quebec99, Quickbeam, Qutezuce, Qxz, R'n'B, R. S. Shaw, RDBrown, RJC, RSido, Ragesoss, Rajwiki123, RandomP, Randomblue, Raul654, Raven in Orbit, Ravidreams, Rdb, Rdsmith4, Red Director, Reddi, Rednblu, Redneckjimmy, Redquark, Retired username, Rettetast, RexNL, Rich Farmbrough, RichG, Richard Durbin's DNA, Ricky81682, Rjwilmsi, Roadnottaken, Robdurbar, RobertG, Rocastelo, RoddyYoung, Rory096, Rotem Dan, Roy Brumback, RoyBoy, RoyLaurie, Royalguard11, RunOrDie, Russ47025, RxS, RyanGerbil10, Ryulong, S77914767, SCEhardt, STAN SWANSON, SWAdair, Sabbre, Safwan40, Sakkura, SallyForth123, Sam Burne James, Samsara, Samuel, Samuel Blanning, SandyGeorgia, Sango123, Sangwine, Savidan, Scarce, Sceptre, Schutz, Sciencechick, Scienceman123, Scincesociety, Sciurinae, Scope creep, Scoterican, Sean William, SeanMack, Seans Potato Business, SebastianHawes, Seldon1, SemperBlotto, Sentausa, Serephine, Shadowlynk, Shanes, ShaunL, Shekharsuman, Shizhao, Shmee47, Shoy, Silsor, SimonD, Sintaku, SirLoin, Sjjupadhyay, Sjollema, Sloth monkey, Slrubenstein, Sly G, SmilesALot, Smithbrenon, Snowmanradio, Snowwolf, Snurks, Solipsist, Someone else, Sonett72, Sopoforic, Spaully, Spectrogram, Spiff, Splette, Spondoolicks, Spongebobsqpants, SpuriousQ, Squidonius, Squirepants101, Statsone, Steel, Steinsky, Stemonitis, Stephenb, SteveHopson, Stevertigo, Stevietheman, Stewartadcock, Stuart7m, Stuhacking, SupaStarGirl, Supspirit, Susvolans, Sverdrup, Swid, Switchercat, Taco325i, Takometer, TakuyaMurata, Tariqabjotu, Tarret, Taulant23, Tavilis, Tazmaniacs, Ted Longstaffe, Tellyaddict, TenOfAllTrades, Terraguy, Test100000, TestPilot, The Rambling Man, The WikiWhippet, TheAlphaWolf, TheChrisD, TheGrza, TheKMan, TheRanger, Thorwald, ThreeDaysGraceFan101, Thue, Tiddly Tom, Tide rolls, TigerShark, TimVickers, Timewatcher, Timir2, Timl2k4, Timrollpickering, Timwi, Tobogganoggin, Toby Bartels, TobyWilson1992, Tom Allen, Tom Harkin's DNA, Tomgally, Toninu, Tony1, Tonyrenplok, Trd300gt, Trent Lott's DNA, Triwbe, Troels Arvin, Tstrobaugh, Tufflaw, Turnstep, Twilight Realm, Ty146Ty146, UBeR, Unint, Unukorno, Usergreatpower, Utcursch, Uthbrian, Vaernnond, Vandelizer, Vanished user, Vary, Virtualphn, Visium, Vividonset, VladimirKorablin, Vsmith, Vyasa, WAS 4.250, WAVEgetarian, WHeimbigner, WJBscribe, Wafulz, WarthogDemon, Wavelength, WelshMatt, West Brom 4ever, Where, Whosasking, Whoutz, Why My Fleece?, Wik, Wiki alf, Wiki emma johnson, Wikiborg, Wikipedia Administration, William Pietri, WillowW, Wimt, Wknight94, Wmahan, Wnt, Wobble, WolfmanSF, Wouterstomp, Wwwwolf, Xy7, YOUR DNA, Yahel Guhan, Yamamoto Ichiro, Yamla, YanWong, Yansa, Yaser al-Nabriss, Yasha, Yomama9753, Younusporteous, Yurik, ZScout370, Zahid Abdassabur, Zahiri, Zanaq, Zazou, Zell Miller's DNA, Zephyris, Zoicon5, Zouavman Le Zouave, Zsinj, Zven, 1329 anonymous edits

Molecular models of DNA *Source:* <http://en.wikipedia.org/w/index.php?oldid=354145216> *Contributors:* Antony-22, Bci2, Chris the speller, CommonsDelinker, David Eppstein, Jwoodger, Michael Hardy, Miym, Nmedard, Oscarthecat, P99am, Tassedethe, The Thing That Should Not Be, 2 anonymous edits

DNA structure *Source:* <http://en.wikipedia.org/w/index.php?oldid=354836243> *Contributors:* Andreww, Antandrus, Bci2, CDN99, Chodges, Cybercobra, DVD R W, DabMachine, David Eppstein, Dysmorodrepanis, Forluvoft, Gene Nygaard, Hannes Röst, Harold f, Joel7687, John Vandenberg, Josq, Leptictidium, Livingrm, Lockesdonkey, Luuva, Maikfr, Mr.Z-man, MrHaiku, Nasz, P99am, Patrick0Moran, Pinethicket, Rajan.kartik, Reinyday, Rich Farmbrough, Simulations, SuperHamster, Thorwald, TimVickers, Tomgally, Wknight94, Yahel Guhan, Yworo, Zephyris, 40 anonymous edits

DNA Dynamics *Source:* <http://en.wikipedia.org/w/index.php?oldid=348694079> *Contributors:* Antony-22, Bci2, Chris the speller, CommonsDelinker, David Eppstein, Jwoodger, Michael Hardy, Miym, Nmedard, Oscarthecat, P99am, Tassedethe, The Thing That Should Not Be, 2 anonymous edits

Interactomics *Source:* <http://en.wikipedia.org/w/index.php?oldid=326148707> *Contributors:* Bci2, Bdevrees, Erick.Antezana, Erodium, J04n, Jong, Jongbhak, Karthik.raman, Lexor, Llull, Niteowlneils, PDH, Pekaje, Rajah, Tucsonnt, 8 anonymous edits

Image Sources, Licenses and Contributors

Image:Fluorescent lighting spectrum peaks labelled.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Fluorescent_lighting_spectrum_peaks_labelled.png *License:* GNU Free Documentation License *Contributors:* H Padleckas, Qef

File:Spectrum of blue flame.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:Spectrum_of_blue_flame.svg *License:* GNU Free Documentation License *Contributors:* user:Deglr6328 on english wikipedia. Recreated from original data

File:Ftir-interferogram.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Ftir-interferogram.png> *License:* GNU Free Documentation License *Contributors:* User:Butenbremer

Image:Interferometer.svg *Source:* <http://en.wikipedia.org/w/index.php?title=File:Interferometer.svg> *License:* GNU Free Documentation License *Contributors:* User:Stannered

Image:HAAtomOrbitals.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:HAAtomOrbitals.png> *License:* GNU Free Documentation License *Contributors:* Admrboltz, Benjah-bmm27, Dbenbenn, Ejdeej, Falcorian, Kborland, MichaelDiederich, Mion, Saperaud, 6 anonymous edits

File:QuantumDot wf.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:QuantumDot_wf.gif *License:* Public Domain *Contributors:* Saumitra R Mehrotra & Gerhard Klimeck

Image:Question_mark2.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:Question_mark2.svg *License:* Public Domain *Contributors:* Original uploader was Acdx at en.wikipedia

File:Rtd seq v3.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Rtd_seq_v3.gif *License:* Public Domain *Contributors:* Saumitra R Mehrotra & Gerhard Klimeck

Image:Bloch sphere.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:Bloch_sphere.svg *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Smite-Meister

Image:Quantum computer.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Quantum_computer.jpg *License:* Creative Commons Attribution 2.5 *Contributors:* Original uploader was Jbw2 at en.wikipedia

Image:BQP complexity class diagram.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:BQP_complexity_class_diagram.svg *License:* Public Domain *Contributors:* User:Mike1024

Image:c60 isosurface.png *Source:* http://en.wikipedia.org/w/index.php?title=File:C60_isosurface.png *License:* GNU Free Documentation License *Contributors:* User:Itamblyn

Image:Calcite.jpg *Source:* <http://en.wikipedia.org/w/index.php?title=File:Calcite.jpg> *License:* unknown *Contributors:* Conscious, Duesentrieb, Lzur, Pieter Kuiper, Ra'ike, Saperaud, Sfu, Wutsje, 5 anonymous edits

Image:Symmetrical stretching.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Symmetrical_stretching.gif *License:* Public Domain *Contributors:* Tiago Becerra Paolini, 1 anonymous edits

Image:Asymmetrical stretching.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Asymmetrical_stretching.gif *License:* Public Domain *Contributors:* Tiago Becerra Paolini

Image:Scissoring.gif *Source:* <http://en.wikipedia.org/w/index.php?title=File:Scissoring.gif> *License:* Public Domain *Contributors:* Tiago Becerra Paolini

Image:Modo rotacao.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Modo_rotacao.gif *License:* Public Domain *Contributors:* Original uploader was Tiago Becerra Paolini at pt.wikipedia

Image:Wagging.gif *Source:* <http://en.wikipedia.org/w/index.php?title=File:Wagging.gif> *License:* Public Domain *Contributors:* Tiago Becerra Paolini

Image:Twisting.gif *Source:* <http://en.wikipedia.org/w/index.php?title=File:Twisting.gif> *License:* Public Domain *Contributors:* Tiago Becerra Paolini

Image:IR spectroscopy apparatus.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:IR_spectroscopy_apparatus.jpg *License:* GNU Free Documentation License *Contributors:* Ewen, Matthias M., Pieter Kuiper

Image:IR summary version 2.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:IR_summary_version_2.gif *License:* unknown *Contributors:* DavidRKelly, Devon Fyson

Image:2dir_pulse_sequence_newversion.png *Source:* http://en.wikipedia.org/w/index.php?title=File:2dir_pulse_sequence_newversion.png *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Pharmacomancer (talk) Original uploader was Pharmacomancer at en.wikipedia

Image:Linearly_pol.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Linearly_pol.png *License:* Public Domain *Contributors:* RASnyder

Image:Circularly_pol.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Circularly_pol.png *License:* Public Domain *Contributors:* RASnyder

Image:Electric Vectors 1.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Electric_Vectors_1.png *License:* unknown *Contributors:* BokicaK

File:Symmetrical stretching.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Symmetrical_stretching.gif *License:* Public Domain *Contributors:* Tiago Becerra Paolini, 1 anonymous edits

File:Asymmetrical stretching.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Asymmetrical_stretching.gif *License:* Public Domain *Contributors:* Tiago Becerra Paolini

File:Scissoring.gif *Source:* <http://en.wikipedia.org/w/index.php?title=File:Scissoring.gif> *License:* Public Domain *Contributors:* Tiago Becerra Paolini

File:Twisting.gif *Source:* <http://en.wikipedia.org/w/index.php?title=File:Twisting.gif> *License:* Public Domain *Contributors:* Tiago Becerra Paolini

File:Wagging.gif *Source:* <http://en.wikipedia.org/w/index.php?title=File:Wagging.gif> *License:* Public Domain *Contributors:* Tiago Becerra Paolini

File:Agitation moléculaire en milieu aqueux.PNG *Source:* http://en.wikipedia.org/w/index.php?title=File:Agitation_moléculaire_en_milieu_aqueux.PNG *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:H'arnet

File:GeneticCode21.svg *Source:* <http://en.wikipedia.org/w/index.php?title=File:GeneticCode21.svg> *License:* unknown *Contributors:* Original uploader was Kosigrim at en.wikipedia

File:Monosodium-glutamate.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Monosodium-glutamate.png> *License:* GNU Free Documentation License *Contributors:* Cacycle, Leyo, Rob Hooft, Samulili

File:H-Gly-Ala-OH.jpg *Source:* <http://en.wikipedia.org/w/index.php?title=File:H-Gly-Ala-OH.jpg> *License:* Creative Commons Attribution-Sharealike 2.5 *Contributors:* Csatazs

File:Zuiterriionball.svg *Source:* <http://en.wikipedia.org/w/index.php?title=File:Zuiterriionball.svg> *License:* Public Domain *Contributors:* user:YassineMrabet

File:Peptidformationball.svg *Source:* <http://en.wikipedia.org/w/index.php?title=File:Peptidformationball.svg> *License:* Public Domain *Contributors:* PatríciaR, Rocket000, YassineMrabet, 2 anonymous edits

File:Aa structure function.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:Aa_structure_function.svg *License:* Public Domain *Contributors:* Jonathan

File:Protein Dynamics Cytochrome C 2NEW smaller.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Protein_Dynamics_Cytochrome_C_2NEW_smaller.gif *License:* GNU Free Documentation License *Contributors:* Original uploader was Zephyris at en.wikipedia

File:Protein-primary-structure.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Protein-primary-structure.png> *License:* Public Domain *Contributors:* National Human Genome Research Institute (NHGRI)

File:Lysozyme crystal1.JPG *Source:* http://en.wikipedia.org/w/index.php?title=File:Lysozyme_crystal1.JPG *License:* Creative Commons Attribution-Sharealike 2.0 *Contributors:* Chrumps, Lode

File:1ezg Tenebrio molitor.png *Source:* http://en.wikipedia.org/w/index.php?title=File:1ezg_Tenebrio_molitor.png *License:* GNU Free Documentation License *Contributors:* Original uploader was WillowW at en.wikipedia

File:Concanavalin A.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Concanavalin_A.png *License:* Public Domain *Contributors:* User:Lijalso

File:Porin.qutemolao.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Porin.qutemolao.png> *License:* Creative Commons Attribution-Sharealike 2.5 *Contributors:* ALoopingIcon

File:Sucrose porin 1a0s.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Sucrose_porin_1a0s.png *License:* GNU Free Documentation License *Contributors:* Opabinia regalis

File:Sucrose specific porin 1A0S.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Sucrose_specific_porin_1A0S.png *License:* GNU Free Documentation License *Contributors:* Snow64

File:StrictosidineSynthase.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:StrictosidineSynthase.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:Hannes Röst

File:Calmodulin 1CLL.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Calmodulin_1CLL.png *License:* Public Domain *Contributors:* Joolz, Lateiner, LeaMaimone, Lode, 1 anonymous edits

File:Haemoglobin-3D-ribbons.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Haemoglobin-3D-ribbons.png> *License:* Public Domain *Contributors:* Benjah-bmm27, Ephemerium, Grafite

File:Hemoglobin t-r state ani.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Hemoglobin_t-r_state_an.gif *License:* GNU Free Documentation License *Contributors:* Conscious, Dbenbenn, Editor at Large, Grafite, Habj, Lennert B, Noca2plus, Tomia, 3 anonymous edits

File:Clostridium perfringens Alpha Toxin Rotate.rsh.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Clostridium_perfringens_Alpha_Toxin_Rotate.rsh.gif *License:* Public Domain *Contributors:* Ramin Herati

File:Bcl-2 Family.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Bcl-2_Family.jpg License: unknown Contributors: Hoffmeier

File:Bcl-2 3D.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Bcl-2_3D.jpg License: GNU Free Documentation License Contributors: CN3D

File:PBB Protein THPO image.jpg Source: http://en.wikipedia.org/w/index.php?title=File:PBB_Protein_THPO_image.jpg License: Public Domain Contributors: www.pdb.org

File:Bence Jones Protein MLE1.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Bence_Jones_Protein_MLE1.jpg License: Public Domain Contributors: Alex McPherson, University of California, Irvine

File:Duck Delta 1 Crystallin.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Duck_Delta_1_Crystallin.jpg License: Public Domain Contributors: Ragesoss

File:Michelson-morley.png Source: <http://en.wikipedia.org/w/index.php?title=File:Michelson-morley.png> License: GNU Free Documentation License Contributors: user:bighead

File:Interferometer.JPG Source: <http://en.wikipedia.org/w/index.php?title=File:Interferometer.JPG> License: Creative Commons Attribution-Sharealike 3.0 Contributors: User:Brews ohare

File:Michelson interferometer schematic.png Source: http://en.wikipedia.org/w/index.php?title=File:Michelson_interferometer_schematic.png License: GNU Free Documentation License Contributors: Teebeutel

File:IR spectroscopy apparatus.jpg Source: http://en.wikipedia.org/w/index.php?title=File:IR_spectroscopy_apparatus.jpg License: GNU Free Documentation License Contributors: Ewen, Matthias M., Pieter Kuiper

File:IR spectrometer.jpg Source: http://en.wikipedia.org/w/index.php?title=File:IR_spectrometer.jpg License: Public Domain Contributors: S.Levchenkov

File:Michelson Interferometer.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Michelson_Interferometer.jpg License: unknown Contributors: Falcorian, Juiced lemon, Teebeutel

File:Ir hcl rot-vib mrtz.svg Source: http://en.wikipedia.org/w/index.php?title=File:Ir_hcl_rot-vib_mrtz.svg License: Creative Commons Attribution-Sharealike 2.5 Contributors: mrtz

File:BandeIR.png Source: <http://en.wikipedia.org/w/index.php?title=File:BandeIR.png> License: GNU Free Documentation License Contributors: User:Grimlock

File:Michelsoninterferometer.jpg Source: <http://en.wikipedia.org/w/index.php?title=File:Michelsoninterferometer.jpg> License: unknown Contributors: Juiced lemon, Skygazer, Tano4595, Teebeutel, Umherirrender, Xorx

File:Michelson couleur.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Michelson_couleur.jpg License: Creative Commons Attribution-Sharealike 2.5 Contributors: NicoB, Teebeutel

File:Bismuthine-2D-IR-MMW-dimensions.png Source: <http://en.wikipedia.org/w/index.php?title=File:Bismuthine-2D-IR-MMW-dimensions.png> License: Public Domain Contributors: Ben Mills

File:Infrared spectrometer.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Infrared_spectrometer.jpg License: GNU Free Documentation License Contributors: ishikawa

Image:Raman energy levels.svg Source: http://en.wikipedia.org/w/index.php?title=File:Raman_energy_levels.svg License: Creative Commons Attribution-Sharealike 3.0 Contributors: User:Moxfyre

Image:CARS diagram.png Source: http://en.wikipedia.org/w/index.php?title=File:CARS_diagram.png License: Creative Commons Attribution-Sharealike 3.0 Contributors: User:Epotma

Image:ev26221 KlyuchevskayaSopka.A2004012.0035.500m.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Ev26221_KlyuchevskayaSopka.A2004012.0035.500m.jpg License: Public Domain Contributors: NASA

Image:FIRST measurement of SF6 and NH3.jpg Source: http://en.wikipedia.org/w/index.php?title=File:FIRST_measurement_of_SF6_and_NH3.jpg License: Creative Commons Attribution-Sharealike 3.0 Contributors: Andre Villemaire

Image:Irspec2.jpg Source: <http://en.wikipedia.org/w/index.php?title=File:Irspec2.jpg> License: Public Domain Contributors: Peterlewis, Qef

Image:Irspec1.jpg Source: <http://en.wikipedia.org/w/index.php?title=File:Irspec1.jpg> License: Public Domain Contributors: Peterlewis

Image:Ozonolysis Scheme.png Source: http://en.wikipedia.org/w/index.php?title=File:Ozonolysis_Scheme.png License: Public Domain Contributors: BenFrantzDale, Benjah-bmm27, ~K

Image:Spec1.jpg Source: <http://en.wikipedia.org/w/index.php?title=File:Spec1.jpg> License: Public Domain Contributors: Bkell, Joey-das-WBF, Peterlewis

Image:Crack2.jpg Source: <http://en.wikipedia.org/w/index.php?title=File:Crack2.jpg> License: Public Domain Contributors: Bkell, Joey-das-WBF, Peterlewis

Image:AminoAcidball.svg Source: <http://en.wikipedia.org/w/index.php?title=File:AminoAcidball.svg> License: Public Domain Contributors: Edgar181, TimVickers, Wutsje, YassineMrabet, 2 anonymous edits

Image:Aa.svg Source: <http://en.wikipedia.org/w/index.php?title=File:Aa.svg> License: Creative Commons Attribution-Sharealike 3.0 Contributors: User:Dancojocari

Image:Lysine fisher struct num.png Source: http://en.wikipedia.org/w/index.php?title=File:Lysine_fisher_struct_num.png License: GNU Free Documentation License Contributors: User:Ppfk

Image:D+L-Alanine.gif Source: <http://en.wikipedia.org/w/index.php?title=File:D+L-Alanine.gif> License: Public Domain Contributors: Ayacop, Edgar181, SharkD, 2 anonymous edits

Image:Amino acid zwitterions.png Source: http://en.wikipedia.org/w/index.php?title=File:Amino_acid_zwitterions.png License: GNU Free Documentation License Contributors: TimVickers

Image:Protein-primary-structure.png Source: <http://en.wikipedia.org/w/index.php?title=File:Protein-primary-structure.png> License: Public Domain Contributors: National Human Genome Research Institute (NHGRI)

Image:L-selenocysteine-2D-skeletal.png Source: <http://en.wikipedia.org/w/index.php?title=File:L-selenocysteine-2D-skeletal.png> License: Public Domain Contributors: Benjah-bmm27

Image:Beta alanine comparison.png Source: http://en.wikipedia.org/w/index.php?title=File:Beta_alanine_comparison.png License: GNU Free Documentation License Contributors: Dr.saptarshi, Edgar181, 1 anonymous edits

Image:Strecker Amino Acid Synthesis Scheme.png Source: http://en.wikipedia.org/w/index.php?title=File:Strecker_Amino_Acid_Synthesis_Scheme.png License: Public Domain Contributors: ~K

Image:Peptidformationball.svg Source: <http://en.wikipedia.org/w/index.php?title=File:Peptidformationball.svg> License: Public Domain Contributors: PatríciaR, Rocket000, YassineMrabet, 2 anonymous edits

Image:Molbio-Header.svg Source: <http://en.wikipedia.org/w/index.php?title=File:Molbio-Header.svg> License: Public Domain Contributors: User:Squidionius

File:Myoglobin.png Source: <http://en.wikipedia.org/w/index.php?title=File:Myoglobin.png> License: Public Domain Contributors: User:AzaToth

File:Peptide group resonance.png Source: http://en.wikipedia.org/w/index.php?title=File:Peptide_group_resonance.png License: GNU Free Documentation License Contributors: Original uploader was WillowW at en.wikipedia

Image:Protein repeating unit.png Source: http://en.wikipedia.org/w/index.php?title=File:Protein_repeating_unit.png License: Public Domain Contributors: User:Edgar181

Image:Peptide bond.png Source: http://en.wikipedia.org/w/index.php?title=File:Peptide_bond.png License: Creative Commons Attribution-Sharealike 3.0 Contributors: webridge

File:Genetic code.svg Source: http://en.wikipedia.org/w/index.php?title=File:Genetic_code.svg License: GNU Free Documentation License Contributors: User:Madprime

File:Proteinviews-1tim.png Source: <http://en.wikipedia.org/w/index.php?title=File:Proteinviews-1tim.png> License: GNU Free Documentation License Contributors: Opabinia regalis

File:Protein Composite.jpg Source: http://en.wikipedia.org/w/index.php?title=File:Protein_Composite.jpg License: Creative Commons Attribution-Sharealike 2.0 Contributors: DENker, GarethWhite, Noca2plus, Photohound, 1 anonymous edits

File:Hexokinase ball and stick model, with substrates to scale copy.png Source: http://en.wikipedia.org/w/index.php?title=File:Hexokinase_ball_and_stick_model_with_substrates_to_scale_copy.png License: Public Domain Contributors: Original uploader was TimVickers at en.wikipedia

File:Mouse-cholera-antibody-1f4x.png Source: <http://en.wikipedia.org/w/index.php?title=File:Mouse-cholera-antibody-1f4x.png> License: GNU Free Documentation License Contributors: Opabinia regalis

File:Localisations02eng.jpg Source: <http://en.wikipedia.org/w/index.php?title=File:Localisations02eng.jpg> License: Creative Commons Attribution-Sharealike 2.5 Contributors: Jeremy Simpson and Rainer Pepperkok

Image:Protein-structure.png Source: <http://en.wikipedia.org/w/index.php?title=File:Protein-structure.png> License: Public Domain Contributors: NHGRI

Image:alpha-h2.png Source: <http://en.wikipedia.org/w/index.php?title=File:Alpha-h2.png> License: Public Domain Contributors: Leonard^Bloom, Monkeybait, Password

Image:alpha-h3.png Source: <http://en.wikipedia.org/w/index.php?title=File:Alpha-h3.png> License: Public Domain Contributors: Leonard^Bloom, Monkeybait, Password

Image:a-amino-acid.png Source: <http://en.wikipedia.org/w/index.php?title=File:A-amino-acid.png> License: Public Domain Contributors: Original uploader was Password at en.wikipedia

Image:amino-CORN.png Source: <http://en.wikipedia.org/w/index.php?title=File:Amino-CORN.png> License: unknown Contributors: Original uploader was Password at en.wikipedia

Image:2-amino-acids.png Source: <http://en.wikipedia.org/w/index.php?title=File:2-amino-acids.png> License: Public Domain Contributors: Original uploader was Password at en.wikipedia

Image:fipsi.png Source: <http://en.wikipedia.org/w/index.php?title=File:Fipsi.png> License: Public Domain Contributors: Monkeybait, Password

Image:Protein folding.png Source: http://en.wikipedia.org/w/index.php?title=File:Protein_folding.png License: Public Domain Contributors: DrKjaergaard, PatríciaR

Image:Protein folding schematic.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Protein_folding_schematic.png *License:* Public Domain *Contributors:* User:Tomixdf

Image:1pkn.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:1pkn.png> *License:* unknown *Contributors:* Original uploader was Dicky2206 at en.wikipedia

File:ADN animation.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:ADN_animation.gif *License:* Public Domain *Contributors:* Aushulz, Bawolff, Brian0918, Elecbullet, Kersti Nebelsiek, Magadan, Mattes, Origamiemensch, Stevenfruitsmaak, Str4nd, Túrelío, 5 anonymous edits

File:Loudspeaker.svg *Source:* <http://en.wikipedia.org/w/index.php?title=File:Loudspeaker.svg> *License:* Public Domain *Contributors:* Bayo, Gmaxwell, Husky, Iamunknown, Nethac DIU, Omegatron, Rocket000, 5 anonymous edits

Image:DNA chemical structure.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:DNA_chemical_structure.svg *License:* Creative Commons Attribution-Sharealike 2.5 *Contributors:* Madprime, Wickey, 1 anonymous edits

File:DNA orbit animated static thumb.png *Source:* http://en.wikipedia.org/w/index.php?title=File:DNA_orbit_animated_static_thumb.png *License:* GNU Free Documentation License *Contributors:* 84user adapting file originally uploaded by Richard Wheeler (Zephyris) at en.wikipedia

Image:GC DNA base pair.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:GC_DNA_base_pair.svg *License:* Public Domain *Contributors:* User:Isilanes

Image:AT DNA base pair.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:AT_DNA_base_pair.svg *License:* Public Domain *Contributors:* User:Isilanes

Image:A-DNA, B-DNA and Z-DNA.png *Source:* [http://en.wikipedia.org/w/index.php?title=File:A-DNA, B-DNA and Z-DNA.png](http://en.wikipedia.org/w/index.php?title=File:A-DNA,_B-DNA_and_Z-DNA.png) *License:* GNU Free Documentation License *Contributors:* Original uploader was Richard Wheeler (Zephyris) at en.wikipedia

Image:Parallel telomere quadruple.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Parallel_telomere_quadruple.png *License:* GNU Free Documentation License *Contributors:* User:Splette

Image:Branch-dna.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Branch-dna.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Original uploader was Peter K. at en.wikipedia

Image:Multi-branch-dna.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Multi-branch-dna.png> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* . Original uploader was Peter K. at en.wikipedia

Image:Cytosine chemical structure.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Cytosine_chemical_structure.png *License:* GNU Free Documentation License *Contributors:* BorisTM, Bryan Derksen, Cacycle, Cwbn (commons), Edgar181, Engineer gena, 1 anonymous edits

Image:5-methylcytosine.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:5-methylcytosine.png> *License:* GNU Free Documentation License *Contributors:* EDUCA33E, Luigi Chiesa, Mysid

Image:Thymine chemical structure.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Thymine_chemical_structure.png *License:* GNU Free Documentation License *Contributors:* Arrowsmaster, BorisTM, Bryan Derksen, Cacycle, Edgar181, Leyo

Image:Benzo[a]pyrene DNA adduct 1JDG.png *Source:* [http://en.wikipedia.org/w/index.php?title=File:Benzo\[a\]pyrene_DNA_adduct_1JDG.png](http://en.wikipedia.org/w/index.php?title=File:Benzo[a]pyrene_DNA_adduct_1JDG.png) *License:* GNU Free Documentation License *Contributors:* Benjah-bmm27, Bstlee, 1 anonymous edits

Image:T7 RNA polymerase at work.png *Source:* http://en.wikipedia.org/w/index.php?title=File:T7_RNA_polymerase_at_work.png *License:* GNU Free Documentation License *Contributors:* User:Splette

Image:DNA replication en.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:DNA_replication_en.svg *License:* Public Domain *Contributors:* user:LadyofHats

Image:Nucleosome 2.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Nucleosome_2.jpg *License:* Public Domain *Contributors:* Original uploader was TimVickers at en.wikipedia

Image:Nucleosome (opposites attracts).JPG *Source:* [http://en.wikipedia.org/w/index.php?title=File:Nucleosome_\(opposites_attracts\).JPG](http://en.wikipedia.org/w/index.php?title=File:Nucleosome_(opposites_attracts).JPG) *License:* Public Domain *Contributors:* Illustration by David S. Goodsell of The Scripps Research Institute (see this site)

Image:Lambda repressor 1LMB.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Lambda_repressor_1LMB.png *License:* GNU Free Documentation License *Contributors:* Original uploader was Zephyris at en.wikipedia

Image:EcoRV 1RVA.png *Source:* http://en.wikipedia.org/w/index.php?title=File:EcoRV_1RVA.png *License:* GNU Free Documentation License *Contributors:* Original uploader was Zephyris at en.wikipedia

Image:Holliday Junction cropped.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Holliday_Junction_cropped.png *License:* GNU Free Documentation License *Contributors:* Original uploader was TimVickers at en.wikipedia

Image:Holliday junction coloured.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Holliday_junction_coloured.png *License:* GNU Free Documentation License *Contributors:* Original uploader was Zephyris at en.wikipedia

Image:Chromosomal Recombination.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:Chromosomal_Recombination.svg *License:* Creative Commons Attribution 2.5 *Contributors:* User:Gringer

Image:DNA nanostructures.png *Source:* http://en.wikipedia.org/w/index.php?title=File:DNA_nanostructures.png *License:* unknown *Contributors:* (Images were kindly provided by Thomas H. LaBean and Hao Yan.)

File:Francis Crick.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Francis_Crick.png *License:* unknown *Contributors:* Photo: Marc Lieberman

File:Rosalind Franklin.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Rosalind_Franklin.jpg *License:* unknown *Contributors:* Avraham, Feydey, Hannes Röst, Shabd sound

File:Raymond Gosling.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Raymond_Gosling.jpg *License:* unknown *Contributors:* Original uploader was Davidruben at en.wikipedia

Image:Spinning DNA.gif *Source:* http://en.wikipedia.org/w/index.php?title=File:Spinning_DNA.gif *License:* Public Domain *Contributors:* USDA

Image:Dna-split2.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Dna-split2.png> *License:* Public Domain *Contributors:* Bobarino, Denniss, Samulili, Wobble

Image:Sarfus.DNABiochip.jpg *Source:* <http://en.wikipedia.org/w/index.php?title=File:Sarfus.DNABiochip.jpg> *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* Nanolane

Image:DNA-(A)80-model.png *Source:* [http://en.wikipedia.org/w/index.php?title=File:DNA-\(A\)80-model.png](http://en.wikipedia.org/w/index.php?title=File:DNA-(A)80-model.png) *License:* Creative Commons Attribution-Sharealike 3.0 *Contributors:* User:P99am

Image:B&Z&A DNA formula.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:B&Z&A_DNA_formula.jpg *License:* Public Domain *Contributors:* Original uploader was Lankenau at en.wikipedia

Image:Circular DNA Supercoiling.png *Source:* http://en.wikipedia.org/w/index.php?title=File:Circular_DNA_Supercoiling.png *License:* GNU Free Documentation License *Contributors:* Richard Wheeler (Zephyris)

Image:Metabolic Network Model for Escherichia coli.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Metabolic_Network_Model_for_Escherichia_coli.jpg *License:* Public Domain *Contributors:* Mdd

License

Creative Commons Attribution-Share Alike 3.0 Unported
<http://creativecommons.org/licenses/by-sa/3.0/>
